AA — AA

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:43:51 ; Search time 32.5 Seconds
(without alignments)
1240.609 Million cell updates/sec

Title: US-09-769-787-162
Perfect score: 1843
Sequence: 1 MSEIKIINAKKIYHDVPVIE.....LNIFSADGSQNLIKGVNHGT 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1843	100.0	364	21	AAY81674	Streptococcus pneu
2	651.5	35.3	381	22	AAU35636	Haemophilus influe
3	651	35.3	373	22	AAB96796	Putative P. abyssi
4	645.5	35.0	371	22	AAB96092	Putative P. abyssi
5	638.5	34.6	352	22	AAU34460	E. coli cellular p
6	627.5	34.0	328	22	AAB88524	Haemophilus influe
7	625.5	33.9	329	22	AAB96235	Putative P. abyssi
8	623.5	33.8	364	22	AAG82000	S. epidermidis ope
9	620.5	33.7	369	22	AAB86541	E. rhapontici PalK
10	617.5	33.5	337	21	AAY81650	Streptococcus pneu
11	616	33.4	467	22	AAU59977	Propionibacterium
12	610	33.1	376	19	AAW60132	M. vaccae pota hom
13	610	33.1	376	20	AAY14879	M. vaccae pota gen
14	605	32.8	386	21	AAY81644	Streptococcus pneu
15	594.5	32.3	408	22	AAG91279	C glutamicum prote
16	592	32.1	332	22	AAG90554	C glutamicum prote
17	592	32.1	376	22	AAG92451	C glutamicum prote
18	591.5	32.1	331	22	AAU36316	Pseudomonas aerugi
19	579	31.4	337	22	ABG24689	Novel human diagno
20	533	28.9	327	19	AAY85963	S. pneumoniae deri
21	506.5	27.5	444	21	AAB05941	Protein deduced fr
22	486.5	26.4	1195	22	ABG24229	Novel human diagno
23	479.5	26.0	313	21	AAY75190	Neisseria gonorrhe
24	478	25.9	464	22	ABG17755	Novel human diagno
25	468.5	25.4	344	22	AAB96082	Putative P. abyssi
26	462	25.1	312	21	AAY75192	Neisseria meningit
27	458	24.9	312	21	AAY75191	Neisseria meningit
28	440.5	23.9	304	22	AAG89989	C glutamicum prote
29	439.5	23.8	458	22	AAU46664	Propionibacterium
30	438.5	23.8	336	22	ABG21669	Novel human diagno
31	424	23.0	159	21	AAB41279	Human ORFX ORF1043
32	422.5	22.9	397	22	AAU35187	Enterococcus faeca
33	421.5	22.9	350	22	AAG90224	C glutamicum prote
34	415	22.5	243	21	AAY81583	Streptococcus pneu
35	415	22.5	423	22	AAU37249	Staphylococcus aur
36	414	22.5	242	18	AAY11137	S. pneumoniae glyc
37	414	22.5	378	22	AAU34303	Staphylococcus aur
38	411.5	22.3	239	22	AAB76771	Corynebacterium gl
39	406	22.0	286	22	AAU59958	Propionibacterium
40	402.5	21.8	375	22	AAG82373	S. epidermidis ope
41	395	21.4	242	22	AAB83175	Corynebacterium th
42	392.5	21.3	382	22	AAU38242	Salmonella typhi c
43	386	20.9	265	22	AAG82703	S. epidermidis ope
44	384.5	20.9	253	21	AAY74487	Neisseria meningit
45	378	20.5	1048	22	ABG22493	Novel human diagno

ALIGNMENTS

RESULT 1
AAY81674
ID AAY81674 standard; Protein; 364 AA.
XX
AC AAY81674;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae protein sequence ID211 - 4127.2.
XX
KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB02451.
XX
PR 27-JUL-1998; 98GB-0016337.
PR 19-MAR-1999; 99US-0125164.
XX
PA (MICR-) MICROBIAL TECHNICS LTD.
XX
PI Gilbert CFG, Hansbro PM;
XX
DR WPI; 2000-195300/17.
XX
PT New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing
PT or inhibiting expression of the protein -

Date no good for us

```
XX
PS   Claim 6; Page 92; 108pp; English.
XX
CC   AAY81501 to AAY81679 represent specifically claimed protein sequences
CC   isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC   specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC   The sequences have antibacterial and antiinflammatory properties.
CC   The protein sequences, and fragments of them, are useful as immunogens
CC   and/or antigens. The nucleotide sequences can be used in vaccines and in
CC   diagnostic assays. The proteins and nucleotides can be useful for the
CC   detection and diagnosis of S. pneumoniae. The protein sequences are also
CC   useful for screening an agent capable of antagonising, inhibiting or
CC   interfering with the function or expression of the proteins in which the
CC   agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC   and meningitis. AAA05591 to AAA05614 represent primers used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   364 AA;


  Query Match            100.0%;  Score 1843;  DB 21;  Length 364;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-167;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mseikiinakkiyhdvpvienlnitipkgslftllgasgcgkttllrmiagfnsieggef 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 yfddtkinnmepskrnigmvfqnyaifphltvrdnvafglmqkkvpkeeliqqtnkylel 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 mqiaqyadrkpdklsggqqqrvtlacalavnpsvllmdeplsnleaklrldmrqaireiq 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 hevgittvyvthdqeeamaisdqiavmkdgviqqigrpkelyhkpanefvatfigrtnii 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 panlekrsdgayivfsdgyalrmpaldqveeqaihvsirpeefikdesgdiegtirdsvy 300

Qy     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 lglntdyfietgfaskiqvseestfeedlqkgnrirlrintqklnifsadgsqnlikgvn 360

Qy     361 HGT 363
           |||
Db     361 hgt 363


RESULT   2
AAU35636
ID   AAU35636 standard; Protein; 381 AA.
XX
AC   AAU35636;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Haemophilus influenzae cellular proliferation protein #277.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Haemophilus influenzae.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS53495.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 11229; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   381 AA;


  Query Match             35.3%;  Score 651.5;  DB 22;  Length 381;
  Best Local Similarity   40.6%;  Pred. No. 1.4e-53;
  Matches  139;  Conservative   70;  Mismatches  106;  Indels   27;  Gaps    7;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:: : || |    :| : |:||  |   |:|| |||||||:||::||   :: |    |
Db      20 ielrsikksygsntiindfnltinngefvtilgpsgcgkttvlrllagleeldsgsiild 79

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             | |:   ||:|  |||:||:|||:|: :||||||  :||| ||:  :  : | ::|:
Db      80 geditnvpaekrhintvfqsyalfphmtifenvafglrmqkvpneeikprvlealrmvql 139

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           : ||||| :||||||||: :| |:   | |||:|| || |: |||  |:| :: :| ::
Db     140 eemadrkptqlsggqqqriaiaravvnkpkvllldeslsaldyklrkqmqqelkmlqrql 199

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           ||| ::|||||||: :|||:| ::: | | | | |:|:|  ||| ||| |||  |:  |
Db     200 gitfifvthdqeeaitmsdrivllrkgkiaqdgspreiyedpanlfvarfigeinvfeat 259

Qy     244 -LEKRSDGAYI---------VFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDE------ 287
           :|::|:   :          :::|     ||     ::| : | :|||: : :|
Db     260 vierkseqvvlanvegricdiytd-----mpv---ekdqklqvllrpedivieeldeneh 311

Qy     288 SGDIEGTIRDSVYLGLNTDYFIETGF-ASKIQVSEESTFEED 328
           |  | | | |  | |:  :  :|      :: |||   | ||
Db     312 skaiighiidrtykgmtlestvefdhngmrvlvse--ffned 351
```

```
RESULT   3
AAB96796
ID   AAB96796 standard; Protein; 373 AA.
XX
AC   AAB96796;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Putative P. abyssi ATPase subunit of ABC transporter #20.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS   Pyrococcus abyssi.
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 7; Pages 1569-1570; 1657pp; French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC   a hyperthermophilic archaeon, which is isolated from deep-sea
CC   hydrothermal vents. The present sequence is one such P. abyssi protein.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence   373 AA;


  Query Match             35.3%;  Score 651;  DB 22;  Length 373;
  Best Local Similarity   38.5%;  Pred. No. 1.5e-53;
  Matches  145;  Conservative   69;  Mismatches  129;  Indels   34;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
             |  |:::  |   | : :    :  ||:||  |    ||| ||||||| ||||||      |:
Db      10 mvevrlenltkkfgnftavnklnltikdgeflvllgpsgcgktttlrmiagleeptegki 69

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           || |  ::   :  | :||| ||||:||::||:|| ||:|| |  || |::|: ::     ||
Db      70 yfgdrevtylpprernismvfqsyavwphmtvydniafplkikkfprdeidkrvrwaael 129

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|| :   || | :|||||:||| :| |: | | ||||||||||:|||||: ||  |:::|
Db     130 lqieelldrypaqlsggqrqrvavaraivvepdvllmdeplsnldaklrvamraeikklq 189

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
            :: :|||:|||||| ||| : |:|||   | : |:| | |:| || : |||||||   :
Db     190 qklkvttiyvthdqveamtmgdriavmnrgqllqvgpptevylkpnsvfvatfigapem- 248

Qy     241 PANLEKRSDGAYIVFSDGYALRMP-----ALDQVEEQAIHVSIRPEEFIKDESGDIEGT- 294
             |: : | |   :   |: : :|      |     : :   ||||         :||
Db     249 --nivevsvgdgylegkgfkielpqdimellrdyigktvlfgirpehm------tvegvs 300

Qy     295 ----IRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQK----------GNRIRLRI 339
               :: :  |    |:    |  : : |     | ::| |           |  : : |
Db     301 elahmkktaklnakvdfvealgtdtilhvk----fgdelvkvklpghipievgkevtivi 356

Qy     340 NTQKLNIFSADGSQNLI 356
           :   :::|  |  : :|
Db     357 dldmmhvfdkdtekaii 373


RESULT   4
AAB96092
ID   AAB96092 standard; Protein; 371 AA.
XX
AC   AAB96092;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Putative P. abyssi ATPase subunit of ABC transporter #5.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS   Pyrococcus abyssi.
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 7; Pages 719-720; 1657pp; French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC   a hyperthermophilic archaeon, which is isolated from deep-sea
CC   hydrothermal vents. The present sequence is one such P. abyssi protein.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence   371 AA;


  Query Match             35.0%;  Score 645.5;  DB 22;  Length 371;
  Best Local Similarity   40.3%;  Pred. No. 4.9e-53;
  Matches  152;  Conservative   66;  Mismatches  116;  Indels   43;  Gaps   12;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |:|:|:||  | : |   :: |:: |  |    ||| ||||||| ||||||      |:
Db       1 maevklinvwkkfgdftavkelsleikdgeflvllgpsgcgktttlrmiagleeptigqi 60

Qy      61 YFDDTKINNME------PSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQT 114
           | ||  : : |       | :|:: ||||:||::||:|| ||:|| |  :||||:|: ::
Db      61 yiddklvadpekgifvppkerdvamvfqsyalyphmtvydniafplklrkvpkqeidrrv 120

Qy     115 NKYLELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQ 174
             :  |:: : :   ||| :||||||||| |  |:   | | ||||||||||:||||: ||
```

```
Db    121 revaemlgltellkrkprelsggqrqrvalgraiirrpkvflmdeplsnldaklrvkmra 180

Qy    175 AIREIQHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFI 234
Db    181 elkklqrqlgvttiyvthdqveamtmgdriavmnrgelqqvgtpdevynkplntfvagfi 240

Qy    235 GR--TNIIPANLEKRSDGAYIVFSDGYALRMPAL-DQVE--------EQAIHVSIRPEE- 282
Db    241 gsppmnfltatvs--sdg----fldfgefrlkllkdqfevleennligkevifgirpedi 294

Qy    283 ----FIKDESGDIEGTIRDSV----YLGLNTDYFIETG---FASKIQVSEESTFEEDLQK 331
Db    295 ydasfirve--epenivrakvdivenlggekivhlrlgdiiftakf--pqesvvee---- 346

Qy    332 GNRIRLRINTQKLNIFS 348
Db    347 gqeaevffdmkkahvfn 363
```

RESULT 5
AAU34460
ID AAU34460 standard; Protein; 352 AA.
XX
AC AAU34460;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #41.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Escherichia coli.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52319.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 10053; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 352 AA;

Query Match 34.6%; Score 638.5; DB 22; Length 352;
Best Local Similarity 39.7%; Pred. No. 2.1e-52;
Matches 136; Conservative 64; Mismatches 116; Indels 27; Gaps 5;

```
Qy      4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
Db     11 velrnvtkrfgsntvidninltipqgqmvtllgpsgcgkttilrlvaglekpsegqifid 70

Qy     64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
Db     71 gedvthrsiqqrdicmvfqsyalfphmslgenvgyglkmlgvpraelkarvkealamvdl 130

Qy    124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
Db    131 egfedrfvdqisggqqqrvalaralilkpkvllfdeplsnldanlrrsmrdkirelqkqf 190

Qy    184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
Db    191 ditslyvthdqseafavsdtvlvmnkghimqigspqdlyrqpasrfmasfmgdanlfpat 250

Qy    244 LEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEG--TIRDSVYL 301
Db    251 f---sdgyvdiy--gyhlprplhfgtqqegm-vgvrpeaitlsdrgeesqrcvirhvaym 304

Qy    302 GLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKL 344
Db    305 gpqyevtve------------------whgqeillqvnatrl 328
```

RESULT 6
AAB88524
ID AAB88524 standard; Protein; 328 AA.
XX
AC AAB88524;
XX
DT 04-JUN-2001 (first entry)
XX
DE Haemophilus influenzae essential bacterial protein SEQ ID NO:66.
XX
KW Haemophilus influenzae; essential bacterial gene; identification;
KW otitis media; meningitis; upper respiratory tract infection;
KW infection; antimicrobial.
XX
OS Haemophilus influenzae.
XX
PN WO200111033-A2.
XX
PD 15-FEB-2001.
XX
PF 03-AUG-2000; 2000WO-US21176.
XX
PR 04-AUG-1999; 99US-0368382.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Chovan LE, Hessler PE, Reich KA;
XX
DR WPI; 2001-147511/15.
DR N-PSDB; AAF94377.
XX
PT Essential bacterial genes from Haemophilus influenzae and methods for
PT identifying 'essential' genes that may be potential therapeutic targets

PT -
XX
PS Claim 9; Page 109-110; 185pp; English.
XX
CC AAF94345 to AAF94409 represent essential bacterial genes from
CC Haemophilus influenzae, which encode the proteins given in AAB88492 to
CC AAB88556. The present invention also describes methods for identifying
CC essential bacterial genes (i.e. those essential to the survival of a
CC bacterium) using a transposition system. The methods are used to
CC identify essential genes from bacteria, especially H. influenzae (which
CC causes otitis media, meningitis and upper respiratory tract infections)
CC which may be used as targets for potential antimicrobial agents.
CC AAF94410 to AAF94416 represent PCR primers used in the exemplification
CC of the present invention.
XX
SQ Sequence 328 AA;

Query Match 34.0%; Score 627.5; DB 22; Length 328;
Best Local Similarity 41.1%; Pred. No. 2.1e-51;
Matches 134; Conservative 58; Mismatches 107; Indels 27; Gaps 4;

```
Qy       8 NAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKI 67
Db      11 nitkafgkavvidnldltikrgtmvtllgpsgcgkttvlrlvaglenptsgqifidgedv 70

Qy      68 NNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYA 127
Db      71 tkssiqnrdicivfqsyalfphmsigdnvgyglkmqgigkeeraqrvkealelvdlagfe 130

Qy     128 DRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITT 187
Db     131 drfvdqisggqqqrvalaralvlkpkvllfdeplsnldanlrrsmrekirelqqrlgits 190

Qy     188 VYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEK- 246
Db     191 lyvthdqteafavsdevivmnkgkimqkapakelylrpnslflanfmgessifdgkleng 250

Qy     247 ----------RSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEE-FIKDESGDIEG-T 294
Db     251 vadingysvplkdaaqfnlpdgecl-------------vgirpeaiylaaegsdaqrce 296

Qy     295 IRDSVYLGLNTDYFIETGFASKIQVS 320
Db     297 iksavymgqplgkwlqtgrgkdllvn 322
```

RESULT 7
AAB96235
ID AAB96235 standard; Protein; 329 AA.
XX
AC AAB96235;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssi ATPase subunit of ABC transporter #7.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT-2000.
XX
PF 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX
DR WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
PS Claim 7; Pages 883-884; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 329 AA;

Query Match 33.9%; Score 625.5; DB 22; Length 329;
Best Local Similarity 41.9%; Pred. No. 3.3e-51;
Matches 140; Conservative 52; Mismatches 91; Indels 51; Gaps 6;

```
Qy      22 LNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEPSKRNIGMVF 81
Db      20 vdlevkdgelisilgpsgcgkttlriiaglerpdkgnvlfdgrdvtflppysrnigfvf 79

Qy      82 QNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRKPDKLSGGQQQR 141
Db      80 qdyalfphmnvfknvafglevrklprkeierkvrevlelvglkgfenrmpeqlsggqqqr 139

Qy     142 VTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITTVYVTHDQEEAMAIS 201
Db     140 valaralviepevllldeplsnldakvreslrmeikriqkelgittmyvthdqeeamais 199

Qy     202 DQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEKRSDGAY-----IVFS 256
Db     200 driavmnfgrieqvgrplelyynpktefvakflgtgnl----lklrstngvaclgelcfn 255

Qy     257 DGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVYLGLNTDYFIETGFASK 316
Db     256 vgv----------dgpvkvffrpesveigkeg-veaev---------mdyeilpg---- 290

Qy     317 IQVSEESTFEEDLQKGNRIRLRINTQKLNIFSAD 350
Db     291 -----------------rirfrlsvqgkeiiaed 307
```

RESULT 8
AAG82000
ID AAG82000 standard; Protein; 364 AA.
XX
AC AAG82000;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:1094.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX

```
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   N-PSDB; AAH52850.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 18; Page 318; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence   364 AA;
```

```
  Query Match             33.8%;  Score 623.5;  DB 22;  Length 364;
  Best Local Similarity   37.9%;  Pred. No. 6e-51;
  Matches  131;  Conservative   72;  Mismatches  118;  Indels   25;  Gaps    4;

Qy       8 NAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKI 67
           :  | : || ::  :|| |  |  :|||| |||||||:|::||||   : |:  : |  |
Db       9 dvskgfedvqilneinidiepgyfytllgpsgcgkttilkliagfeypdsgdiiykdkpi 68

Qy      68 NNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYA 127
             | |:|| :  |||:||:|||| | ||:|:||  ||: | |: ::  : |:|:::: |
Db      69 gkmppnkrkvntvfqdyalfphlnvfdniayglklkklskseikrkvtealqlvklsgye 128

Qy     128 DRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITT 187
             |:   :||||:||| :| |:   | :||:|| || |: ||| :|:  :||:|  :|||
Db     129 hrqiqgmsgggkqrvaiaraivnepeillldeslsaldlklrtemqyllrelqsrlgitf 188

Qy     188 VYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEKR 247
           ::|||||||:|:|| | ||||| ||| | | ::| :| | ||| ||| :||:   : :
Db     189 ifvthdqeealalsdyifvmkdgkiqqfgtpidiydepvnrfvadfigesnivhgtmve- 247

Qy     248 SDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEE--FIKDESGDIEGTIRDSVYLGLNT 305
             ::|   |      :   | : : | ||||:   :    |  :  :   :: |::
Db     248 ---dfvvniygqnfdcvdmgikenkkvevvirpedislvsqndglfkakvdsmlfrgvhy 304

Qy     306 DYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           :          |   :   :|  :|         :|:| |: |  |
Db     305 e----------icckdrkgyewviq---------stkkanvgsevg 331
```

```
RESULT   9
AAB86541
ID   AAB86541 standard; Protein; 369 AA.
XX
AC   AAB86541;
XX
DT   05-NOV-2001  (first entry)
XX
DE   E. rhapontici PalK protein SEQ ID 21.
XX
KW   Palatinase; regulatory region; cell-active promoter; trehalulase;
KW   selection marker; plant; ATP-binding protein; palK.
XX
OS   Erwinia rhapontici.
XX
PN   DE10045182-A1.
XX
PD   16-AUG-2001.
XX
PF   13-SEP-2000; 2000DE-1045182.
XX
PR   14-FEB-2000; 2000DE-1006463.
XX
PA   (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
XX
PI   Boernke FF,  Sonnewald U;
XX
DR   WPI; 2001-497706/55.
DR   N-PSDB; AAH48815, AAH48816.
XX
PT   New recombinant nucleic acids encoding a palatinase or a trehalulase
PT   and also comprising plant transcription regulatory sequences are useful
PT   as selection markers for transformed plants  -
XX
PS   Disclosure; Page 42-43; 60pp; German.
XX
CC   This invention describes a novel recombinant nucleic acid molecule (I),
CC   comprising regulatory sequences of a plant cell-active promoter,
CC   operatively linked to a DNA sequence encoding a protein with palatinase
CC   and/or trehalulase activity, and to plant cell regulatory sequences for
CC   transcription termination and/or polyadenylation signal. The invention is
CC   used as a selection marker, particularly for transformed plant cells or
CC   microorganisms. This sequence represents the Erwinia rhapontici
CC   ATP-binding protein PalK which is described in the method of the
CC   invention.
XX
SQ   Sequence   369 AA;
```

```
  Query Match             33.7%;  Score 620.5;  DB 22;  Length 369;
  Best Local Similarity   36.8%;  Pred. No. 1.2e-50;
  Matches  135;  Conservative   76;  Mismatches  131;  Indels   25;  Gaps    8;

Qy       1 MSEIKIINAKKIY-HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
           |::: :   :| |     || :||: |  |    ::| |||||:|||||||||   | ||
Db       1 maqltldkiqkrygakaevirslnlqiksgefvvivgpsgcgkstllrmiagleeisggg 60

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
             | |    |:  |::| ||||||:||::||::|  |:|| |      | |: |:  :
Db      61 myidghyanddspaergigmvfqsyalyphmsvyqnmafalelahcskaeidqrvrecar 120

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           ::|:    :|:|  |||||:||| :  |:   | : | |||||||:| ||: ||  :  :
Db     121 ilqlepllerrpkdlsgggrqrvaigraiireprlflfdeplsnldaslrvqmrmevsal 180

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RT 237
             :|:| :|||||| ||| ::|:| |:  | |:|:| | ||| :|||||| |||  :
Db     181 hkrlgvtiiyvthdqveamtladrivvlnqggnieqvgtplelydqpanefvaqfigspkm 240

Qy     238 NIIPANLEKRSDGAYIVFSDG-----YALRMPALDQVEEQAIHVSIRPEEFIKDESGDIE 292
           |:||| | :  :   :|  |       ::  ||  : | ::::: ||||      ||::|
Db     241 nlipatlrrsgeqqsvveldngktlvlsiatpa--eaegrsvnigirpehi---rsgnve 295

Qy     293 -----GTIRDSVYLGLNTDYFIETGFASK---IQVSEESTFEEDLQKGNRIRLRINTQKL 344
                | :     :|  |  ::: | | :   :: :| |     :  |  : :|:  :
Db     296 qceyqgevmfveqmgnetllyldngnagepwvvrhaersa----ihvgqtvgvrlpvecc 351

Qy     345 NIFSADG 351
             :| : |
```

```
Db    352 ylfdshg 358


RESULT  10
AAY81650
ID   AAY81650 standard; Protein; 337 AA.
XX
AC   AAY81650;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Streptococcus pneumoniae type 4 protein sequence #150.
XX
KW   Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW   antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW   pneumococcal disease.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO200006737-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-GB02451.
XX
PR   27-JUL-1998;   98GB-0016337.
PR   19-MAR-1999;   99US-0125164.
XX
PA   (MICR-) MICROBIAL TECHNICS LTD.
XX
PI   Gilbert CFG,  Hansbro PM;
XX
DR   WPI; 2000-195300/17.
XX
PT   New Streptococcal protein, useful as a vaccine, for diagnosis of
PT   pneumococcal diseases and for screening agents capable of antagonizing
PT   or inhibiting expression of the protein  -
XX
PS   Claim 1; Page 87; 108pp; English.
XX
CC   AAY81501 to AAY81679 represent specifically claimed protein sequences
CC   isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC   specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC   The sequences have antibacterial and antiinflammatory properties.
CC   The protein sequences, and fragments of them, are useful as immunogens
CC   and/or antigens. The nucleotide sequences can be used in vaccines and in
CC   diagnostic assays. The proteins and nucleotides can be useful for the
CC   detection and diagnosis of S. pneumoniae. The protein sequences are also
CC   useful for screening an agent capable of antagonising, inhibiting or
CC   interfering with the function or expression of the proteins in which the
CC   agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC   and meningitis. AAA05591 to AAA05614 represent primers used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   337 AA;


  Query Match             33.5%;  Score 617.5;  DB 21;  Length 337;
  Best Local Similarity   39.3%;  Pred. No. 2e-50;
  Matches  143;  Conservative   61;  Mismatches  107;  Indels   53;  Gaps    8;

Qy      4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
          ||  | :  | |   |:|||: | :|  || || |||||:| || : ||     |    :
Db      2 ikfdniqikygdfvaidnlnldihegefftflgpsgcgksttlralvgfldpssgsievn 61

Qy     64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
           | : ::|| || ||:|||:||:|| :|| ||:|||| ||| : :  : :     ::|
Db     62 gtdvthlepekrgigivfqsyalfptmtvfdniafglkvkkvapdvikakvsavaakiki 121

Qy    124 A-QYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHE 182
          : |   |   :|||||||| || || : | :| :|||||||:||||:|:|: :: :| |
Db    122 sdqqlqrnvselsggqqqrvalaralvlepkilcldeplsnldaklrvdlrkelkrlqke 181

Qy    183 VGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPA 242
          :||||:|||||||||: :||:|||  :| |:|:| | |:||    |||  |||  |::
Db    182 lgittlyvthdqeealtlsdriavfnngyieqvgtpveiyhnsqtefvcdfigdinvltd 241

Qy    243 N-----LEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGD--IEGTI 295
              | | :     ||             :|::  :: :    | ::   |  ::|||
Db    242 etvhevllknts----vf------------ledkkgyirlekvrfnreteqdfilkgti 284

Qy    296 RDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNL 355
           |  : |:   |         |:|||                   :| ||: | | ||
Db    285 idvefsgvtihy--------tikvse------------------sqilnvtsid-sqaa 316

Qy    356 IKGV 359
          |: |
Db    317 irsv 320


RESULT  11
AAU59977
ID   AAU59977 standard; Protein; 467 AA.
XX
AC   AAU59977;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #20873.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-616774/71.
DR   N-PSDB; AAS59607.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris -
XX
PS   Claim 3; SEQ ID No 21172; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
```

```
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   467 AA;


  Query Match             33.4%;  Score 616;  DB 22;  Length 467;
  Best Local Similarity   35.2%;  Pred. No. 4.5e-50;
  Matches  135;  Conservative   73;  Mismatches  128;  Indels   48;  Gaps    8;

Qy       1 MSEIKIINAKKIY--HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGG 58
Db     102 matvqykeatrvypgtdraavskldldigdgefmvlvgpsgcgkstslrmlagleevnag 161

Qy      59 EFYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYL 118
Db     162 siwigdrdvtdlppkdrdiamvfqnyalyphmtvadnmgfalkmqnvpkaerqqrvleaa 221

Qy     119 ELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIRE 178
Db     222 kllgledflnrkpkalsgggrqrvamgraivrnpqvflmdeplsnldaklrvqtrtqiaa 281

Qy     179 IQHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--- 235
Db     282 lqqrlgvttvyvthdqveamtmgdrvavmrdgvlqqvdsplalydtpknlfvagfigspa 341

Qy     236 ----RTNIIPANLEKRSDGAYIV-FSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGD 290
Db     342 mnlmegdvvdggvel---gdyvvpvsrdvlakvp-----netkltlgirpeaftlaseg- 392

Qy     291 IEGTIRDSVYLGLNTDYFIETG-----FASKIQVSEESTFEE--------------DLQK 331
Db     393 ----------igldvavveelgadaylygslvgsgkqasiedgeahqvvarissrrppqr 442

Qy     332 GNRIRLRINTQKLNIFSADGSQNL 355
Db     443 geqvrlgvdpasvhvfsqetteri 466


RESULT  12
AAW60132
ID   AAW60132 standard; Protein; 376 AA.
XX
AC   AAW60132;
XX
DT   25-AUG-1998  (first entry)
XX
DE   M. vaccae pota homologue protein sequence.
XX
KW   Mycobacterium vaccae; antigen; therapy; prevention; cytokine production;
KW   M. avium; M. tuberculosis; immune response enhancer; cell proliferation;
KW   mycobacteria infection; vaccine; cancer; pota gene.
XX
OS   Mycobacterium vaccae.
XX
PN   WO9808542-A2.
XX
PD   05-MAR-1998.
XX
PF   28-AUG-1997;   97WO-NZ000105.
XX
PR   12-JUN-1997;   97US-0873970.
PR   29-AUG-1996;   96US-0705347.
XX
PA   (GENE-) GENESIS RES & DEV CORP.
XX
PI   Hiyama J,  Prestidge RL,  Scott LM,  Skinner MA,  Tan P;
PI   Visser E;
XX
DR   WPI; 1998-216926/19.
DR   N-PSDB; AAV34595.
XX
PT   Mycobacterium vaccae polypeptides - used to develop products for use
PT   in detection, therapy and prevention of mycobacteria infections or
PT   as immune response enhancers
XX
PS   Example 8; Pages 103-104; 153pp; English.
XX
CC   This represents a Mycobacterium vaccae pota homologue that is homologous
CC   to E. coli pota protein. The invention provides M. vaccae polypeptides
CC   that comprise an immunogenic portion of a soluble M. vaccae antigen, or
CC   a variant, where the antigen induces an immune response in patients
CC   previously exposed to a mycobacterium. Such M. vaccae polypeptides can
CC   be used in methods for enhancing non-specific immune response. The
CC   methods and products can be used for the detection, treatment and
CC   prevention of infectious diseases caused by mycobacteria such as
CC   M. vaccae, M. avium or M. tuberculosis. The products also have the
CC   ability to induce cell proliferation and cytokine production (e.g.
CC   interferon-gamma and interleukin-12 production) in T cells, NK cells,
CC   B cells, or macrophages. They can be used for enhancing immune
CC   responses for use in vaccines or immunotherapy of infectious diseases
CC   and cancers.
XX
SQ   Sequence   376 AA;


  Query Match             33.1%;  Score 610;  DB 19;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-49;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
Db       2 ieidhvtkrfgdylavadadfsiapgeffsmlgpsgcgktttlrmiagfetptegairle 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
Db      62 gadvsrtppnkrnvntvfqhyalfphmtvwdnvaygprskklgkgevrkrvdelleivrl 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
Db     122 tefaerrpaqlsggqqqrvalaralvnypsaslldeplgaldlklrhvmqfelkriqrev 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
Db     182 gitfiyvthdqeealtmsdriavmnagnveqigspteiydrpatvfvasfigqanlwagr 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
Db     242 ctgrsnrdyveidvlgstlkarpgettiepgghatlmvrperirvtpgsqdaptgdvacv 301

Qy     292 EGTIRDSVYLG 302
Db     302 ratvtdltfqg 312


RESULT  13
AAY14879
ID   AAY14879 standard; protein; 376 AA.
XX
AC   AAY14879;
XX
DT   25-OCT-1999  (first entry)
XX
DE   M. vaccae pota gene homologue amino acid sequence.
XX
KW   Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
KW   dendritic cell maturation; infectious disease; immune disorder; cancer;
KW   respiratory system; mycobacterial infection; allergy; tuberculosis;
KW   leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KW   dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KW   squamous cell carcinoma; melanoma.
XX
OS   Mycobacterium vaccae.
```

XX
PN WO9932634-A2.
XX
PD 01-JUL-1999.
XX
PF 23-DEC-1998; 98WO-NZ00189.
XX
PR 04-DEC-1998; 98US-0205426.
PR 23-DEC-1997; 97US-0996624.
PR 23-DEC-1997; 97US-0997080.
PR 23-DEC-1997; 97US-0997362.
PR 11-JUN-1998; 98US-0095855.
PR 17-SEP-1998; 98US-0156181.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
XX
DR WPI; 1999-430163/36.
DR N-PSDB; AAZ11330.
XX
PT Enhancing immune response to an antigen
XX
PS Example 15; Page 179; 243pp; English.
XX
CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
CC T cells and natural killer cells, to stimulate the production of
CC cytokines, to enhance the expression of co-stimulatory molecules on
CC dendritic cells and monocytes, and to enhance dendritic cell maturation
CC and function. The proteins can be expressed by standard recombinant
CC methodology. Pharmaceutical compositions comprising the proteins or
CC nucleic acid sequences encoding the proteins can be used for the
CC treatment, prevention, and detection of disorders including infectious
CC diseases, immune disorders and cancer. In particular, the compounds and
CC methods are used for treatment of diseases of the respiratory system,
CC such as mycobacterial infections, asthma, allergies, tuberculosis,
CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma.
XX
SQ Sequence 376 AA;

Query Match 33.1%; Score 610; DB 20; Length 376;
Best Local Similarity 42.4%; Pred. No. 1.2e-49;
Matches 132; Conservative 61; Mismatches 106; Indels 12; Gaps 5;

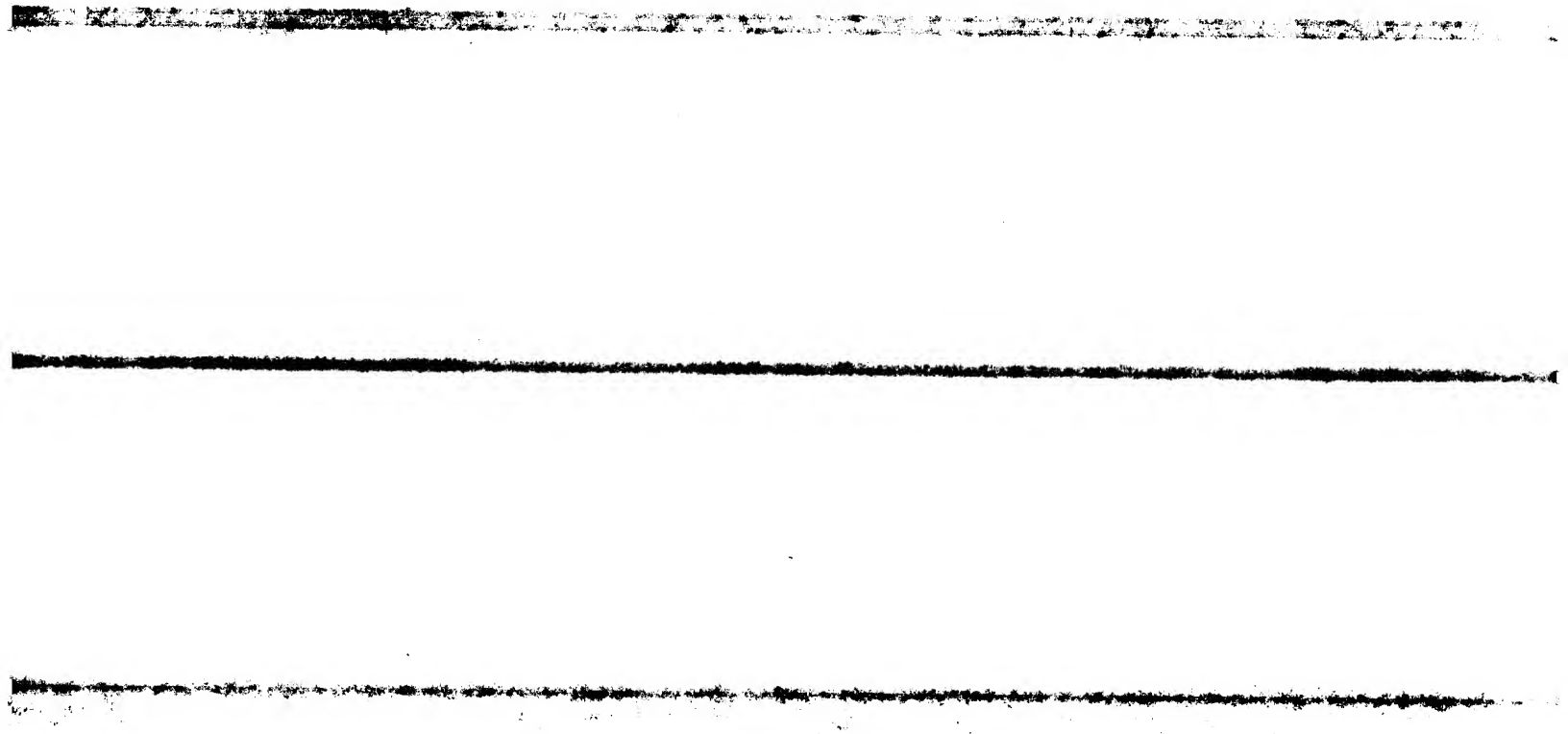
```
Qy      4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
          |:| :  | : |   : : : :|  |  |::|| ||||||| ||||||| :   |    :
Db      2 ieidhvtkrfgdylavadadfsiapgeffsmlgpsgcgktttlrmiagfetptegairle 61

Qy     64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
            ::   |:|||:  |||:||:|||:|| ||||:|   ||: | |: :: :: ||:::: 
Db     62 gadvsrtppnkrnvntvfqhyalfphmtvwdnvaygprskklgkgevrkrvdelleivrl 121

Qy    124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           ::|:|:| :|||||||| || ||   || ||:|||| |: |||  |:  :: || ||
Db    122 tefaerrpaqlsggqqqrvalaralvnypsallldeplgaldlklrhvmqfelkriqrev 181

Qy    184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
          ||| :||||||||: :||:|||| | ::||| | |:| :||  |||:||: |:
Db    182 gitfiyvthdqeealtmsdriavmnagnveqigspteiydrpatvfvasfigqanlwagr 241

Qy    244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
             ||:  |:     |  |:    :   |   | :  :|||        :|  :||   :
Db    242 ctgrsnrdyveidvlgstlkarpgettiepgghatlmvrperirvtpgsqdaptgdvacv 301

Qy    292 EGTIRDSVYLG 302
            |: |  : |
Db    302 ratvtdltfqg 312
```

RESULT 14
AAY81644
ID AAY81644 standard; Protein; 386 AA.
XX
AC AAY81644;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae type 4 protein sequence #144.
XX
KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB02451.
XX
PR 27-JUL-1998; 98GB-0016337.
PR 19-MAR-1999; 99US-0125164.
XX
PA (MICR-) MICROBIAL TECHNICS LTD.
XX
PI Gilbert CFG, Hansbro PM;
XX
DR WPI; 2000-195300/17.
XX
PT New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing
PT or inhibiting expression of the protein -
XX
PS Claim 1; Page 87; 108pp; English.
XX
CC AAY81501 to AAY81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties.
CC The protein sequences, and fragments of them, are useful as immunogens
CC and/or antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AAA05591 to AAA05614 represent primers used in the
CC exemplification of the present invention.
XX
SQ Sequence 386 AA;

Query Match 32.8%; Score 605; DB 21; Length 386;
Best Local Similarity 41.7%; Pred. No. 3.8e-49;
Matches 128; Conservative 67; Mismatches 100; Indels 12; Gaps 5;

```
Qy      4 IKIINAKKIYHD--VPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFY 61
          |:  |  |:: |    |::::|  : :|  :|||||| ||:|:| :|||      |:
Db      6 iefknvskvfedsntkvlkdinfeleegkfytllgasgsgkstilniiaglldattgdim 65

Qy     62 FDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELM 121
           |  :||::  :||::  |||:||:|||: | :|||| |  :|: |:|: |:  : |:::
Db     66 ldgvrindiptnkrdvhtvfqsyalfphmnvfenvafplrlrkidkkeieqrvaevlkmv 125

Qy    122 QIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQH 181
          |:  |  |  ||||||:||| :| |:   | |:|:|||| |: ||| ||:  :||:|
Db    126 qlegyekrsirklsggqrqrvaiaraiinqprvvlldeplsaldlklrtdmqyelrelqq 185

Qy    182 EVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIP 241
          :||| |:||||||||:|:|| | || || | | | | ::| :| | |||||| :||:|
```

```
Db    186 rlgitfvfvthdqeealamsdwifvmndgeivqsgtpvdiydepinhfvatfigesnilp 245

Qy    242 ANLEKRSDGAYIVFSDGYALRMPALD--QVEEQAIHVSIRPEE--FIKDESGDIEGTIRD 297
          : :     |:|  :|   |  |:|      : : | ||||:      | | ::  :
Db    246 gtmie----dylvefng--krfeavdggmkpnepvevvirpedlritlpeegklqvkvdt 299

Qy    298 SVYLGLN 304
          :: |::
Db    300 qlfrgvh 306


RESULT  15
AAG91279
ID   AAG91279 standard; Protein; 408 AA.
XX
AC   AAG91279;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 5033.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   N-PSDB; AAH66498.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
XX
PS   Claim 17; SEQ ID NO: 5033; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   408 AA;


  Query Match             32.3%;  Score 594.5;  DB 22;  Length 408;
  Best Local Similarity   36.5%;  Pred. No. 4.2e-48;
  Matches  139;  Conservative   62;  Mismatches  133;  Indels   47;  Gaps    8;

Qy     17 PVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEPSKRN 76
          | :: ||: |  |    |:| |||||:| |||:||   |: |    |      : |  |:
Db     19 pavdklnleiadgeflvlvgpsgcgkstslrmlaglepidegrllidgkdatelrpqdrd 78

Qy     77 IGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRKPDKLSG 136
          | ||||:||::|::|||||: | |  :|| | |: ::  :   ::|:  | ||||  |||
Db     79 iamvfqsyalypnmtvrdnmgfalknqkvakaeiekrvaeasrilqldpyldrkpaalsg 138

Qy    137 GQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITTVYVTHDQEE 196
          ||:||| :  |:   ||| ||||||||:||||:  |  |  :|  :|:|||||||| |
Db    139 gqrqrvamgraivrepsvfcmdeplsnldaklrvstraeisglqrrmgvttvyvthdqve 198

Qy    197 AMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTNIIPANLEKRSDGAYIV 254
          || : |::||:  ||:||:  |: ||  ||| |||:|||    |:|   :     |  :
Db    199 amtmgdrvavlllgvlqqvdtpqnlydypanafvasfigspsmnliegtir----gdkvt 254

Qy    255 FSDGYALRMP---------ALDQVEEQAIHVSIRPEEF---IKDESGDIEGTIRDSVYLG 302
             |  : :|           |: | : : |  |||       :||| : | :     ||
Db    255 lgtgiqisvpdevaaevrnnpdrfegrpvivgarpehmylttanesgavlgevshidelg 314

Qy    303 L--------------NTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQ------ 342
                          |||   | |    ::|:     | |     : || |  :
Db    315 adsmvyvlasgvknpntdllge-gipedmrvtvvgaeetd-----karlgirverhhglk 368

Qy    343 ---KLNIFSADGSQNLIKGVN 360
             |::: :|    :|  |::
Db    369 agdkvhvvaapkdvhlfdgld 389


Search completed: July 30, 2002, 15:47:43
Job time: 232 sec
```



GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:46:16 ; Search time 13.08 Seconds
(without alignments)
677.866 Million cell updates/sec

Title: US-09-769-787-162
Perfect score: 1843
Sequence: 1 MSEIKIINAKKIYHDVPVIE.....LNIFSADGSQNLIKGVNHGT 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	610	33.1	376	2	US-08-997-080-89	Sequence 89, Appl
2	610	33.1	376	2	US-08-997-362-89	Sequence 89, Appl
3	610	33.1	376	3	US-08-873-970-89	Sequence 89, Appl
4	610	33.1	376	4	US-09-095-855-89	Sequence 89, Appl
5	610	33.1	376	4	US-09-324-542-89	Sequence 89, Appl
6	349	18.9	242	2	US-08-402-804-6	Sequence 6, Appli
7	345	18.7	244	4	US-08-919-573-2	Sequence 2, Appli
8	345	18.7	244	4	US-08-919-573-4	Sequence 4, Appli
9	338.5	18.4	203	4	US-08-858-207A-315	Sequence 315, App
10	293.5	15.9	111	1	US-08-466-886-36	Sequence 36, Appl
11	293.5	15.9	111	4	US-08-469-617-36	Sequence 36, Appl
12	271	14.7	265	2	US-08-898-779-2	Sequence 2, Appli
13	271	14.7	265	2	US-08-898-779-4	Sequence 4, Appli
14	271	14.7	265	4	US-09-224-502-2	Sequence 2, Appli
15	271	14.7	265	4	US-09-224-502-4	Sequence 4, Appli
16	260.5	14.1	1375	3	US-08-665-259-26	Sequence 26, Appl
17	260.5	14.1	1375	3	US-08-762-500-26	Sequence 26, Appl
18	259	14.1	262	4	US-08-961-083-30	Sequence 30, Appl
19	252.5	13.7	1275	3	US-09-120-513-2	Sequence 2, Appli
20	252.5	13.7	1275	4	US-09-450-105-2	Sequence 2, Appli
21	248.5	13.5	1279	2	US-08-784-649A-2	Sequence 2, Appli
22	248.5	13.5	1280	2	US-08-583-276-19	Sequence 19, Appl
23	248.5	13.5	1280	6	5206352-4	Patent No. 5206352
24	246.5	13.4	202	4	US-08-858-207A-332	Sequence 332, App
25	246.5	13.4	1684	3	US-08-665-259-25	Sequence 25, Appl
26	246.5	13.4	1684	3	US-08-762-500-25	Sequence 25, Appl
27	246.5	13.4	1704	3	US-08-762-500-75	Sequence 75, Appl
28	243.5	13.2	1280	2	US-08-752-447-2	Sequence 2, Appli
29	236.5	12.8	1457	3	US-08-665-259-27	Sequence 27, Appl
30	236.5	12.8	1457	3	US-08-762-500-27	Sequence 27, Appl
31	235.5	12.8	1308	2	US-08-996-644-2	Sequence 2, Appli
32	235.5	12.8	1308	3	US-09-352-552-2	Sequence 2, Appli
33	235	12.8	233	4	US-08-993-825-2	Sequence 2, Appli
34	234.5	12.7	1334	2	US-08-996-545-2	Sequence 2, Appli
35	234.5	12.7	1334	4	US-09-328-320-2	Sequence 2, Appli
36	234	12.7	707	3	US-08-772-270A-4	Sequence 4, Appli
37	232	12.6	233	4	US-09-161-662-2	Sequence 2, Appli
38	231.5	12.6	1408	1	US-08-612-521-2	Sequence 2, Appli
39	231	12.5	1349	2	US-08-612-734B-2	Sequence 2, Appli
40	230.5	12.5	246	4	US-09-627-376-15	Sequence 15, Appl
41	230.5	12.5	1307	1	US-08-395-246C-2	Sequence 2, Appli
42	230	12.5	79	2	US-08-997-080-87	Sequence 87, Appl
43	230	12.5	79	2	US-08-997-362-87	Sequence 87, Appl
44	230	12.5	79	3	US-08-873-970-87	Sequence 87, Appl
45	230	12.5	79	4	US-09-095-855-87	Sequence 87, Appl

ALIGNMENTS

```
RESULT   1
US-08-997-080-89
; Sequence 89, Application US/08997080
; Patent No. 5968524
;  GENERAL INFORMATION:
;    APPLICANT:  WATSON, JAMES D.
;    APPLICANT:  TAN, PAUL L.J.
;    TITLE OF INVENTION:  METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,080
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  376 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-997-080-89


  Query Match             33.1%;  Score 610;  DB 2;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 2.6e-55;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;
```

```
Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
Db       2 IEIDHVTKRFGDYLAVADADFSIAPGEFFSMLGPSGCGKTTTLRMIAGFETPTEGAIRLE 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
Db      62 GADVSRTPPNKRNVNTVFQHYALFPHMTVWDNVAYGPRSKKLGKGEVRKRVDELLEIVRL 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
Db     122 TEFAERRPAQLSGGQQQRVALARALVNYPSALLLDEPLGALDLKLRHVMQFELKRIQREV 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
Db     182 GITFIYVTHDQEEALTMSDRIAVMNAGNVEQIGSPTEIYDRPATVFVASFIGQANLWAGR 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
Db     242 CTGRSNRDYVEIDVLGSTLKARPGETTIEPGGHATLMVRPERIRVTPGSQDAPTGDVACV 301

Qy     292 EGTIRDSVYLG 302
Db     302 RATVTDLTFQG 312


RESULT   2
US-08-997-362-89
; Sequence 89, Application US/08997362
; Patent No. 5985287
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  194
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,362
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/873,970
;      FILING DATE:  June 12, 1997
;      APPLICATION NUMBER:  U.S. Patent Application No. 5985287 08/705,347
;      FILING DATE:  August 29, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  376 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-997-362-89


  Query Match             33.1%;  Score 610;  DB 2;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 2.6e-55;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
Db       2 IEIDHVTKRFGDYLAVADADFSIAPGEFFSMLGPSGCGKTTTLRMIAGFETPTEGAIRLE 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
Db      62 GADVSRTPPNKRNVNTVFQHYALFPHMTVWDNVAYGPRSKKLGKGEVRKRVDELLEIVRL 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
Db     122 TEFAERRPAQLSGGQQQRVALARALVNYPSALLLDEPLGALDLKLRHVMQFELKRIQREV 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
Db     182 GITFIYVTHDQEEALTMSDRIAVMNAGNVEQIGSPTEIYDRPATVFVASFIGQANLWAGR 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
Db     242 CTGRSNRDYVEIDVLGSTLKARPGETTIEPGGHATLMVRPERIRVTPGSQDAPTGDVACV 301

Qy     292 EGTIRDSVYLG 302
Db     302 RATVTDLTFQG 312


RESULT   3
US-08-873-970-89
; Sequence 89, Application US/08873970
; Patent No. 6001361
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Hiyama, Jun
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Scott, Linda
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/873,970
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
```

```
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  376 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-873-970-89


  Query Match             33.1%;  Score 610;  DB 3;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 2.6e-55;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:| :  | : |   : : : :|  |  |::|| ||||||| ||||||| :   |    :
Db       2 IEIDHVTKRFGDYLAVADADFSIAPGEFFSMLGPSGCGKTTTLRMIAGFETPTEGAIRLE 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             ::   |:|||:  |||:||:|||:|| ||||:|   ||: | |: :: :: ||:::: 
Db      62 GADVSRTPPNKRNVNTVFQHYALFPHMTVWDNVAYGPRSKKLGKGEVRKRVDELLEIVRL 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           ::|:|:| :|||||||| || ||   || ||:|||  |: |||  |:  :: || ||
Db     122 TEFAERRPAQLSGGQQQRVALARALVNYPSALLLDEPLGALDLKLRHVMQFELKRIQREV 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           ||| :|||||||||: :||:|||  | ::||| | |:| :||  |||:|||: |:
Db     182 GITFIYVTHDQEEALTMSDRIAVMNAGNVEQIGSPTEIYDRPATVFVASFIGQANLWAGR 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
             ||:  |:     |  |:   :  |   | :  :|||       :|  :||   :
Db     242 CTGRSNRDYVEIDVLGSTLKARPGETTIEPGGHATLMVRPERIRVTPGSQDAPTGDVACV 301

Qy     292 EGTIRDSVYLG 302
             |: |  : |
Db     302 RATVTDLTFQG 312


RESULT   4
US-09-095-855-89
; Sequence 89, Application US/09095855
; Patent No. 6160093
;  GENERAL INFORMATION:
;    APPLICANT:  Tan, Paul
;    APPLICANT:  Visser, Elizabeth
;    APPLICANT:  Skinner, Margot
;    APPLICANT:  Prestidge, Ross
;    TITLE OF INVENTION:  Compounds and Methods for
;    TITLE OF INVENTION:  Treatment and Diagnosis of Mycobacterial Infections
;    NUMBER OF SEQUENCES:  208
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Offices of Ann W. Speckman
;      STREET:  2601 Elliott Avenue, Suite 4185
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/095,855
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/705,347
;      FILING DATE:  29-AUG-1996
;      APPLICATION NUMBER:  08/873,970
;      FILING DATE:  12-JUN-1997
;      APPLICATION NUMBER:  08/997,362
;      FILING DATE:  23-DEC-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sleath, Janet
;      REGISTRATION NUMBER:  37,007
;      REFERENCE/DOCKET NUMBER:  11000.1002c3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-269-0565
;      TELEFAX:  206-269-0563
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  376 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-095-855-89


  Query Match             33.1%;  Score 610;  DB 4;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 2.6e-55;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:| :  | : |   : : : :|  |  |::|| ||||||| ||||||| :   |    :
Db       2 IEIDHVTKRFGDYLAVADADFSIAPGEFFSMLGPSGCGKTTTLRMIAGFETPTEGAIRLE 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             ::   |:|||:  |||:||:|||:|| ||||:|   ||: | |: :: :: ||:::: 
Db      62 GADVSRTPPNKRNVNTVFQHYALFPHMTVWDNVAYGPRSKKLGKGEVRKRVDELLEIVRL 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           ::|:|:| :|||||||| || ||   || ||:|||  |: |||  |:  :: || ||
Db     122 TEFAERRPAQLSGGQQQRVALARALVNYPSALLLDEPLGALDLKLRHVMQFELKRIQREV 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           ||| :|||||||||: :||:|||  | ::||| | |:| :||  |||:|||: |:
Db     182 GITFIYVTHDQEEALTMSDRIAVMNAGNVEQIGSPTEIYDRPATVFVASFIGQANLWAGR 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
             ||:  |:     |  |:   :  |   | :  :|||       :|  :||   :
Db     242 CTGRSNRDYVEIDVLGSTLKARPGETTIEPGGHATLMVRPERIRVTPGSQDAPTGDVACV 301

Qy     292 EGTIRDSVYLG 302
             |: |  : |
Db     302 RATVTDLTFQG 312


RESULT   5
US-09-324-542-89
; Sequence 89, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Tan, Paul L.J.
;  APPLICANT: Prestidge, Ross
;  TITLE OF INVENTION: Methods and Compounds for the Treatment
;  TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
;  FILE REFERENCE: 11000.1007c1
;  CURRENT APPLICATION NUMBER: US/09/324,542
;  CURRENT FILING DATE: 1999-06-02
;  EARLIER APPLICATION NUMBER: US 08/997,080
;  EARLIER FILING DATE: 1997-12-23
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 89
;   LENGTH: 376
;   TYPE: PRT
;   ORGANISM: Mycobacterium vaccae
US-09-324-542-89
```

```
  Query Match             33.1%;  Score 610;  DB 4;  Length 376;
  Best Local Similarity   42.4%;  Pred. No. 2.6e-55;
  Matches  132;  Conservative   61;  Mismatches  106;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:| :  | : |   : : : :|  |  |::|| |||||| ||||||| :  |    :
Db       2 IEIDHVTKRFGDYLAVADADFSIAPGEFFSMLGPSGCGKTTTLRMIAGFETPTEGAIRLE 61

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             ::   |:|||:  |||:||:|||:|| ||||:|   ||: | |: :: :: ||::::
Db      62 GADVSRTPPNKRNVNTVFQHYALFPHMTVWDNVAYGPRSKKLGKGEVRKRVDELLEIVRL 121

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           ::|:|:| :|||||||||| || ||   || ||:|||| |: |||  |:  :: || ||
Db     122 TEFAERRPAQLSGGQQQRVALARALVNYPSALLLDEPLGALDLKLRHVMQFELKRIQREV 181

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           ||| :|||||||||: :||:|||  | ::||| | |:| :||  |||:|||: |:
Db     182 GITFIYVTHDQEEALTMSDRIAVMNAGNVEQIGSPTEIYDRPATVFVASFIGQANLWAGR 241

Qy     244 LEKRSDGAYIVFSD-GYALRMPALDQVEEQAIHVS--IRPEEF-----IKD-ESGD---I 291
             ||:  |:     |  |:     :   |   | :  :|||        :|  :||   :
Db     242 CTGRSNRDYVEIDVLGSTLKARPGETTIEPGGHATLMVRPERIRVTPGSQDAPTGDVACV 301

Qy     292 EGTIRDSVYLG 302
             |: |  : |
Db     302 RATVTDLTFQG 312


RESULT   6
US-08-402-804-6
; Sequence 6, Application US/08402804
; Patent No. 5874300
;  GENERAL INFORMATION:
;    APPLICANT:  Blaser, Martin J.
;    APPLICANT:  Pei, Zhiheng
;    TITLE OF INVENTION:  Campylobacter Jejuni Antigens, And
;    TITLE OF INVENTION:  Methods For Their Production And Use
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OSTROLENK, FABER, GERB & SOFFEN
;      STREET:  1180 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-8403
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/402,804
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/114,420
;      FILING DATE:  30-AUG-1993
;      APPLICATION NUMBER:  US 08/112,387
;      FILING DATE:  27-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/986,928
;      FILING DATE:  08-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/612,330
;      FILING DATE:  13-NOV-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gray III, William O.
;      REGISTRATION NUMBER:  30,944
;      REFERENCE/DOCKET NUMBER:  P/1261-13
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 382-0700
;      TELEFAX:  (212) 382-0888
;      TELEX:  236925
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  242 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-402-804-6


  Query Match             18.9%;  Score 349;  DB 2;  Length 242;
  Best Local Similarity   31.6%;  Pred. No. 2.2e-28;
  Matches   75;  Conservative   55;  Mismatches  103;  Indels    4;  Gaps    2;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:: |  | |    |::  |::: :|    ::| || ||:| :| : |   :  ||   :
Db       2 IELKNVNKYYGTHHVLKIFNLSVKEGEKLVIIGPSGSGKSTTIRCMNGLEEVSSGEVVVN 61

Qy      64 DTKINN---MEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQ-KKVPKEELIQQTNKYLE 119
             :  :|:   :|  ::   ||||:: ::||:|| |:    |: :|  |:|  :   |||:
Db      62 NLVLNHKNKIEICRKYCAMVFQHFNLYPHMTVLQNLTLAPMKLQKKSKKEAEETAFKYLK 121

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :: :   |:  |  |||||||||| :| :|      :| ||| | |: :   ::   ::||
Db     122 VVGLLDKANVYPATLSGGQQQRVAIARSLCTKKPYILFDEPTSALDPETIQEVLDVMKEI 181

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR 236
           |:   | | |||:   |  ::|:|  |:|| | :   | | :  |  |    |:|:
Db     182 SHQSNTTMVVVTHEMGFAKEVADRIIFMEDGAIVEENIPSEFFSNPKTERARLFLGK 238


RESULT   7
US-08-919-573-2
; Sequence 2, Application US/08919573
; Patent No. 6346392
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    TITLE OF INVENTION:  NOVEL GLUTAMINE TRANSPORT
;    TITLE OF INVENTION:  ATP-BINDING PROTEIN
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/919,573
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50597
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  244 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
US-08-919-573-2


  Query Match             18.7%;  Score 345;  DB 4;  Length 244;
  Best Local Similarity   31.5%;  Pred. No. 5.7e-28;
  Matches   75;  Conservative   59;  Mismatches   98;  Indels    6;  Gaps    3;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |||  |   | :    |:: :|: | :| :  ::| || ||:|||| :        |:  |:
Db       6 IKIENLHKSFGKNEVLKGINLEIKRGEVVVIIGPSGSGKSTLLRSMNLLEEATKGKVIFE 65

Qy      64 DTKI----NNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQ-QTNKYL 118
             |    |::    :  :|||| : :|||::|| :|:    :: |    |: : :  | |
Db      66 GVDITDKKNDLFAMREKMGMVFQQFNLFPNMTVIENITLSPIKTKGESREVAEKRAQKLL 125

Qy     119 ELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIRE 178
           | : :   ||  |  ||||||||: :|  ||: | ||| ||| | |: ::  ::   :::
Db     126 EKVGLPDKADAYPQSLSGGQQQRIAIARGLAMEPDVLLFDEPTSALDPEMVGEVLAVMQD 185

Qy     179 IQHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR 236
           :  : |:| | |||:   |  ::|::  | |||: : |:|:::: :   :    |: :
Db     186 LA-KSGMTMVIVTHEMGFAREVADRVIFMADGVVVEDGTPEQIFEQTQEQRTKDFLSK 242


RESULT   8
US-08-919-573-4
; Sequence 4, Application US/08919573
; Patent No. 6346392
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    TITLE OF INVENTION:  NOVEL GLUTAMINE TRANSPORT
;    TITLE OF INVENTION:  ATP-BINDING PROTEIN
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/919,573
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50597
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  244 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-919-573-4


  Query Match             18.7%;  Score 345;  DB 4;  Length 244;
  Best Local Similarity   31.5%;  Pred. No. 5.7e-28;
  Matches   75;  Conservative   59;  Mismatches   98;  Indels    6;  Gaps    3;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |||  |   | :    |:: :|: | :| :  ::| || ||:|||| :        |:  |:
Db       6 IKIENLHKSFGKNEVLKGINLEIKRGEVVVIIGPSGSGKSTLLRSMNLLEEATKGKVIFE 65

Qy      64 DTKI----NNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQ-QTNKYL 118
             |    |::    :  :|||| : :|||::|| :|:    :: |    |: : :  | |
Db      66 GVDITDKKNDLFAMREKMGMVFQQFNLFPNMTVIENITLSPIKTKGESREVAEKRAQKLL 125

Qy     119 ELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIRE 178
           | : :   ||  |  ||||||||: :|  ||: | ||| ||| | |: ::  ::   :::
Db     126 EKVGLPDKADAYPQSLSGGQQQRIAIARGLAMEPDVLLFDEPTSALDPEMVGEVLAVMQD 185

Qy     179 IQHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR 236
           :  : |:| | |||:   |  ::|::  | |||: : |:|:::: :   :    |: :
Db     186 LA-KSGMTMVIVTHEMGFAREVADRVIFMADGVVVEDGTPEQIFEQTQEQRTKDFLSK 242


RESULT   9
US-08-858-207A-315
; Sequence 315, Application US/08858207A
; Patent No. 6348328
;  GENERAL INFORMATION:
;    APPLICANT:  Black, Michael
;    APPLICANT:  Hodgson, John
;    APPLICANT:  Knowles, David
;    APPLICANT:  Nicholas, Richard
;    APPLICANT:  Stodola, Robert
;    TITLE OF INVENTION:  No. 6348328el Compounds
;    NUMBER OF SEQUENCES:  552
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/858,207A
;      FILING DATE:  09-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/017670
;      FILING DATE:  14-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gimmi, Edward R
;      REGISTRATION NUMBER:  38,891
;      REFERENCE/DOCKET NUMBER:  P50475
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-4478
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  315:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6348328e
US-08-858-207A-315


  Query Match             18.4%;  Score 338.5;  DB 4;  Length 203;
  Best Local Similarity   35.5%;  Pred. No. 2e-27;
  Matches   72;  Conservative   57;  Mismatches   69;  Indels    5;  Gaps    4;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           | |:: ||  |:: |  :: : :::||:  :  ::| ||  |||||||:||  :|: |:
```

```
Db      1 MLELRNIN--KVFGDKQILSDFSLSIPEKQILAIVGPSGGSKTTLLRMLAGLETIDSGQI 58

Qy     61 YFDDTKINNMEPSKRN-IGMVFQNYAIFPHLTVRDNVAFGLMQKK-VPKEELIQQTNKYL 118
          :::   :   |  ||| :| |||:: :||||:| :|:    ::   : :||  :: :  |
Db     59 FYNGQPLELDELQKRNLLGFVFQDFQLFPHLSVLENLTLSPVKTMGMKQEEAEKKASGLL 118

Qy    119 ELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIRE 178
          | : :  :|:  |  |||||:||| || |: ::| ::  ||| | |: :|||:: : | :
Db    119 EQLGLGGHAESYPFSLSGGQKQRVALARAMMIDPEIIGYDEPTSALDPELRLEVEKLILQ 178

Qy    179 IQHEVGITTVYVTHDQEEAMAIS 201
             |:|:| : |||| : |  |:
Db    179 -NRELGMTQIVVTHDLQFAENIA 200


RESULT  10
US-08-466-886-36
; Sequence 36, Application US/08466886
; Patent No. 5776677
;  GENERAL INFORMATION:
;    APPLICANT:  Tsui, Lap-Chee
;    APPLICANT:  Riordan, John R.
;    APPLICANT:  Rommens, Johanna M.
;    APPLICANT:  Kerem, Bat-Sheva
;    APPLICANT:  Collins, Francis S.
;    APPLICANT:  Iannuzzi, Michael C.
;    APPLICANT:  Drumm, Mitchell L.
;    APPLICANT:  Buckwald, Manuel
;    TITLE OF INVENTION:  Cystic Fibrosis Gene
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,886
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldstein, Jorge A.
;      REGISTRATION NUMBER:  29,021
;      REFERENCE/DOCKET NUMBER:  1329.0010006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-466-886-36


  Query Match             15.9%;  Score 293.5;  DB 1;  Length 111;
  Best Local Similarity   40.0%;  Pred. No. 3.8e-23;
  Matches   64;  Conservative   22;  Mismatches   25;  Indels   49;  Gaps    2;

Qy     11 KIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNM 70
          | : :| | :::|| | :|     :| |||||:||||||||
Db      1 KAWGEVVVSKDINIDIHEGEFVVFVGPSGCGKSTLLRMIAG------------------- 41

Qy     71 EPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRK 130
                :|||||:||::|||:| :|::|||                          : ||||
Db     42 ------VGMVFQSYALYPHLSVAENMSFGL-----------------------KPADRK 71

Qy    131 PDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRL 170
          |  ||||:|||| :   |   ||| |:|||||||:| ||:
Db     72 PKALSGGRQQRVAIGRTLVAEPSVFLLDEPLSNLDAALRV 111


RESULT  11
US-08-469-617-36
; Sequence 36, Application US/08469617
; Patent No. 6201107
;  GENERAL INFORMATION:
;    APPLICANT:  Tsui, Lap-Chee
;    APPLICANT:  Riordan, John R.
;    APPLICANT:  Rommens, Johanna M.
;    APPLICANT:  Kerem, Bat-Sheva
;    APPLICANT:  Collins, Francis S.
;    APPLICANT:  Iannuzzi, Michael C.
;    APPLICANT:  Drumm, Mitchell L.
;    APPLICANT:  Buckwald, Manuel
;    TITLE OF INVENTION:  Cystic Fibrosis Gene
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,617
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Goldstein, Jorge A.
;      REGISTRATION NUMBER:  29,021
;      REFERENCE/DOCKET NUMBER:  1329.0010008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-469-617-36


  Query Match             15.9%;  Score 293.5;  DB 4;  Length 111;
  Best Local Similarity   40.0%;  Pred. No. 3.8e-23;
  Matches   64;  Conservative   22;  Mismatches   25;  Indels   49;  Gaps    2;

Qy     11 KIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNM 70
          | : :| | :::|| | :|     :| |||||:||||||||
Db      1 KAWGEVVVSKDINIDIHEGEFVVFVGPSGCGKSTLLRMIAG------------------- 41

Qy     71 EPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRK 130
                :|||||:||::|||:| :|::|||                          : ||||
Db     42 ------VGMVFQSYALYPHLSVAENMSFGL-----------------------KPADRK 71

Qy    131 PDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRL 170
          |  ||||:|||| :   |   ||| |:|||||||:| ||:
Db     72 PKALSGGRQQRVAIGRTLVAEPSVFLLDEPLSNLDAALRV 111


RESULT  12
```

```
US-08-898-779-2
; Sequence 2, Application US/08898779
; Patent No. 5882891
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    APPLICANT:  Lonetto, Michael Arthur
;    APPLICANT:  Warren, Patrick Vernon
;    TITLE OF INVENTION:  NOVEL Ferrichrome transport
;    TITLE OF INVENTION:  ATP-Binding Protein
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/898,779
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50444-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  265 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-898-779-2


  Query Match             14.7%;  Score 271;  DB 2;  Length 265;
  Best Local Similarity   28.8%;  Pred. No. 3.4e-20;
  Matches   62;  Conservative   56;  Mismatches   91;  Indels    6;  Gaps    3;

Qy      13 YHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEP 72
           | |  :|  |:: || | : :::| :|||:|||: ::   ::: || : |   |:
Db      13 YGDNTIINKLDVEIPDGKVTSIIGPNGCGKSTLLKALSRLLAVKEGEVFLDGENIHTQST 72

Qy      73 SK--RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE-LMQIA---QY 126
           :  : | :: |:  :   ||| : |::|    :     |  :  | ::  |::    :
Db      73 KEIAKKIAILPQSPEVADGLTVGELVSYGRFPHQKGFGRLTAEDKKEIDWAMEVTGTDTF 132

Qy     127 ADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGIT 186
             |  : |||||:||| :| |||    :: :||| : |:   :|:: : :::: | | |
Db     133 RHRSINDLSGGQRQRVWIAMALAQRTDIIFLDEPTTYLDICHQLEILELVQKLNQEQGCT 192

Qy     187 TVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKEL 221
             | | ||  :|:  || :  ||:| |   |  :::
Db     193 IVMVLHDINQAIRFSDHLIAMKEGDIIATGSTEDV 227


RESULT  13
US-08-898-779-4
; Sequence 4, Application US/08898779
; Patent No. 5882891
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    APPLICANT:  Lonetto, Michael Arthur
;    APPLICANT:  Warren, Patrick Vernon
;    TITLE OF INVENTION:  NOVEL Ferrichrome transport
;    TITLE OF INVENTION:  ATP-Binding Protein
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/898,779
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50444-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  265 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-898-779-4


  Query Match             14.7%;  Score 271;  DB 2;  Length 265;
  Best Local Similarity   28.8%;  Pred. No. 3.4e-20;
  Matches   62;  Conservative   56;  Mismatches   91;  Indels    6;  Gaps    3;

Qy      13 YHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEP 72
           | |  :|  |:: || | : :::| :|||:|||: ::   ::: || : |   |:
Db      13 YGDNTIINKLDVEIPDGKVTSIIGPNGCGKSTLLKALSRLLAVKEGEVFLDGENIHTQST 72

Qy      73 SK--RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE-LMQIA---QY 126
           :  : | :: |:  :   ||| : |::|    :     |  :  | ::  |::    :
Db      73 KEIAKKIAILPQSPEVADGLTVGELVSYGRFPHQKGFGRLTAEDKKEIDWAMEVTGTDTF 132

Qy     127 ADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGIT 186
             |  : |||||:||| :| |||    :: :||| : |:   :|:: : :::: | | |
Db     133 RHRSINDLSGGQRQRVWIAMALAQRTDIIFLDEPTTYLDICHQLEILELVQKLNQEQGCT 192

Qy     187 TVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKEL 221
             | | ||  :|:  || :  ||:| |   |  :::
Db     193 IVMVLHDINQAIRFSDHLIAMKEGDIIATGSTEDV 227


RESULT  14
US-09-224-502-2
; Sequence 2, Application US/09224502
; Patent No. 6264955
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    APPLICANT:  Lonetto, Michael Arthur
;    APPLICANT:  Warren, Patrick Vernon
;    TITLE OF INVENTION:  NOVEL Ferrichrome transport
;    TITLE OF INVENTION:  ATP-Binding Protein
;    NUMBER OF SEQUENCES:  4
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/224,502
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/898,779
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50444-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  265 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-224-502-2


  Query Match             14.7%;  Score 271;  DB 4;  Length 265;
  Best Local Similarity   28.8%;  Pred. No. 3.4e-20;
  Matches   62;  Conservative   56;  Mismatches   91;  Indels    6;  Gaps    3;

Qy      13 YHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEP 72
           | |  :|  |:: || | : :::| :|||| :|||: ::   ::: || : |   |:
Db      13 YGDNTIINKLDVEIPDGKVTSIIGPNGCGKSTLLKALSRLLAVKEGEVFLDGENIHTQST 72

Qy      73 SK--RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE-LMQIA---QY 126
           :  : | :: |:  :   ||| : |::|    :     |  :  | ::  |::     :
Db      73 KEIAKKIAILPQSPEVADGLTVGELVSYGRFPHQKGFGRLTAEDKKEIDWAMEVTGTDTF 132

Qy     127 ADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGIT 186
             | : |||||:||| :| |||    :: :||| : |:   :|:: : ::::  | | |
Db     133 RHRSINDLSGGQRQRVWIAMALAQRTDIIFLDEPTTYLDICHQLEILELVQKLNQEQGCT 192

Qy     187 TVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKEL 221
           | | ||  :|:  || :  ||:| |   |  :::
Db     193 IVMVLHDINQAIRFSDHLIAMKEGDIIATGSTEDV 227


RESULT  15
US-09-224-502-4
; Sequence 4, Application US/09224502
; Patent No. 6264955
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin, Karl Russel
;    APPLICANT:  Lonetto, Michael Arthur
;    APPLICANT:  Warren, Patrick Vernon
;    TITLE OF INVENTION:  NOVEL Ferrichrome transport
;    TITLE OF INVENTION:  ATP-Binding Protein
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  997 Lenox Drive, Building 3, Suite 210
;      CITY:  Lawrenceville
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08543
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/224,502
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/898,779
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bloom, Allen
;      REGISTRATION NUMBER:  29,135
;      REFERENCE/DOCKET NUMBER:  P50444-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-520-3214
;      TELEFAX:  609-520-3259
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  265 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-224-502-4


  Query Match             14.7%;  Score 271;  DB 4;  Length 265;
  Best Local Similarity   28.8%;  Pred. No. 3.4e-20;
  Matches   62;  Conservative   56;  Mismatches   91;  Indels    6;  Gaps    3;

Qy      13 YHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEP 72
           | |  :|  |:: || | : :::| :|||| :|||: ::   ::: || : |   |:
Db      13 YGDNTIINKLDVEIPDGKVTSIIGPNGCGKSTLLKALSRLLAVKEGEVFLDGENIHTQST 72

Qy      73 SK--RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE-LMQIA---QY 126
           :  : | :: |:  :   ||| : |::|    :     |  :  | ::  |::     :
Db      73 KEIAKKIAILPQSPEVADGLTVGELVSYGRFPHQKGFGRLTAEDKKEIDWAMEVTGTDTF 132

Qy     127 ADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGIT 186
             | : |||||:||| :| |||    :: :||| : |:   :|:: : ::::  | | |
Db     133 RHRSINDLSGGQRQRVWIAMALAQRTDIIFLDEPTTYLDICHQLEILELVQKLNQEQGCT 192

Qy     187 TVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKEL 221
           | | ||  :|:  || :  ||:| |   |  :::
Db     193 IVMVLHDINQAIRFSDHLIAMKEGDIIATGSTEDV 227


Search completed: July 30, 2002, 15:48:30
Job time: 134 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:45:46 ; Search time 20.92 Seconds
(without alignments)
1667.324 Million cell updates/sec

Title: US-09-769-787-162
Perfect score: 1843
Sequence: 1 MSEIKIINAKKIYHDVPVIE.....LNIFSADGSQNLIKGVNHGT 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1843	100.0	363	2	D95028	ABC transporter, A
2	1817	98.6	363	2	F97899	hypothetical prote
3	731	39.7	365	2	AF3119	hypothetical prote
4	731	39.7	374	2	A98168	hypothetical prote
5	714.5	38.8	355	2	A72279	sugar ABC transpor
6	701	38.0	368	2	A72261	spermidine/putresc
7	671	36.4	352	2	H97617	ABC transporter, A
8	671	36.4	352	2	AF2840	hypothetical prote
9	670.5	36.4	358	2	H72555	probable transport
10	668	36.2	373	2	A82398	maltose/maltodextr
11	666.5	36.2	346	2	B98169	ATP-binding transp
12	666.5	36.2	346	2	AE3118	hypothetical prote
13	662	35.9	373	2	AH3583	glucose ABC transp
14	661.5	35.9	351	2	AD3587	SN-glycerol-3-phos
15	661.5	35.9	397	2	T35802	probable polyamine
16	661	35.9	375	2	H71241	probable multiple
17	658.5	35.7	392	2	T45204	probable sugar ABC
18	657.5	35.7	369	2	H72272	sugar ABC transpor
19	656	35.6	364	2	H89879	hypothetical prote
20	653	35.4	358	2	E95842	probable lactose t
21	652	35.4	353	2	AB3650	iron(III)-transpor
22	651.5	35.3	381	2	B64118	spermidine/putresc
23	651	35.3	364	2	D75187	probable maltose/m
24	650.5	35.3	347	2	A70180	spermidine/putresc
25	650	35.3	373	2	C71220	probable sugar-bin
26	649.5	35.2	357	2	E71123	probable sugar tra
27	649	35.2	378	2	D83784	sugar ABC transpor
28	648	35.2	381	2	F72756	probable multiple
29	645.5	35.0	371	2	B75207	multiple sugar-bin
30	644.5	35.0	329	2	G81344	ABC transport syst
31	644	34.9	378	2	A40840	spermidine/putresc
32	644	34.9	378	2	C90825	spermidine/putresc
33	644	34.9	378	2	F85683	spermidine/putresc
34	643.5	34.9	370	2	A83621	probable ATP-bindi
35	643.5	34.9	381	2	D87637	hypothetical prote
36	642.5	34.9	408	2	S77405	hypothetical prote
37	641.5	34.8	353	2	D98334	ATP-binding compon
38	641.5	34.8	353	2	AF2948	hypothetical prote
39	641.5	34.8	377	2	F82201	spermidine/putresc
40	639.5	34.7	352	2	E90680	probable ABC trans
41	639.5	34.7	353	2	AF3580	iron(III)-transpor
42	638.5	34.6	352	2	F64751	probable ABC-type
43	638	34.6	363	2	S34734	ATP-binding protei
44	637.5	34.6	362	2	E95972	probable sugar upt
45	637	34.6	355	2	H96012	probable sugar upt

ALIGNMENTS

RESULT 1
D95028
ABC transporter, ATP-binding protein SP0242 [imported] - Streptococcus pneumoniae (st
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: D95028
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: D95028
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-363 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK74421.1; PID:g14971711; GSPDB:GN00164; TIGR
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0242

Query Match 100.0%; Score 1843; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.9e-122;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300

Qy     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360

Qy     361 HGT 363
```

```
          |||
Db    361 HGT 363


RESULT   2
F97899
hypothetical protein ABC-NBD [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C;Accession: F97899
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: F97899
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-363 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99026.1; PID:g15457768; GSPDB:GN00174
C;Genetics:
A;Gene: ABC-NBD


  Query Match             98.6%;  Score 1817;  DB 2;  Length 363;
  Best Local Similarity   98.3%;  Pred. No. 2e-120;
  Matches  357;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
          |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||: |||
Db      1 MSEIEIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIKDGEF 60

Qy     61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
          ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db     61 YFDDTKINNMEPSKRNIGLVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120

Qy    121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQTIREIQ 180

Qy    181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240

Qy    241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300

Qy    301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360
          |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db    301 LGLNTDYFIETGFASKIQVSEESTFEEELQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360

Qy    361 HGT 363
          |||
Db    361 HGT 363


RESULT   3
AF3119
hypothetical protein Atu4578 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AF3119
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AF3119
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45372.1; PID:g17743068; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4578
A;Map position: linear chromosome


  Query Match             39.7%;  Score 731;  DB 2;  Length 365;
  Best Local Similarity   45.1%;  Pred. No. 5e-44;
  Matches  156;  Conservative   66;  Mismatches   98;  Indels   26;  Gaps    6;

Qy      1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
          |: ::| : || | ||  : ::||:|| || ||||| |||||||||| ||||:  : |
Db      5 MAAVEITSIKKSYRDVVALSDINISIPSGSFFTLLGPSGCGKTTLLRTIAGFHQQDSGSI 64

Qy     61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
            :   | ::   ||::|||||:||:|||::| ||:|||| |:|    |: ::  | |::
Db     65 SIERQAIEHVPAHKRDVGMVFQDYAVFPHISVFDNIAFGLKQRKSSSAEIRERVGKILDV 124

Qy    121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
          :|:| || | | :|||||||| || || :|| ||||||||||:||||:|:|: :||||
Db    125 VQLAPYAKRMPHELSGGQQQRVGLARALVINPKVLLMDEPLSNLDAKLRVDLRRELRREIQ 184

Qy    181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
          : |||||||||||||:|:|| : ||  |||||   | |:|: ||| |||:|:|  | :
Db    185 QAMNITTVYVTHDQEEALAMSDLVCVMYGGVIQQAAPPWEVYNNPANRFVASFVGANNFL 244

Qy    241 PANLEKRSDGAYIVFSDGYALRMPALDQV-EEQAIHVSIRPEEFI-------KDESGDIE 292
          :   |: |   : |    : :|  | : ::: |  :||||           ||  ::
Db    245 TLD---RTGGQTSISS--LKVTLPQADAISKDRTIVAAIRPEAISIGPSIGNCDERIELP 299

Qy    293 GTIRDSVYLG--LNTDYFIETGFASK----------IQVSEESTF 325
          |||   : |  :|    : :|   :           :| ::::||
Db    300 VTIRQVSFTGREMNVAAVLSSGEEIEAITKPSPEIIALQPNQKTTF 345


RESULT   4
A98168
hypothetical protein AGR_L_589 [imported] - Agrobacterium tumefaciens (strain C58, Ce
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: A98168
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: A98168
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88867.1; PID:g15158633; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_589
A;Map position: linear chromosome


  Query Match             39.7%;  Score 731;  DB 2;  Length 374;
  Best Local Similarity   45.1%;  Pred. No. 5.2e-44;
  Matches  156;  Conservative   66;  Mismatches   98;  Indels   26;  Gaps    6;

Qy      1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
          |: ::| : || | ||  : ::||:|| || ||||| |||||||||| ||||:  : |
Db     14 MAAVEITSIKKSYRDVVALSDINISIPSGSFFTLLGPSGCGKTTLLRTIAGFHQQDSGSI 73

Qy     61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
            :   | ::   ||::|||||:||:|||::| ||:|||| |:|    |: ::  | |::
Db     74 SIERQAIEHVPAHKRDVGMVFQDYAVFPHISVFDNIAFGLKQRKSSSAEIRERVGKILDV 133
```

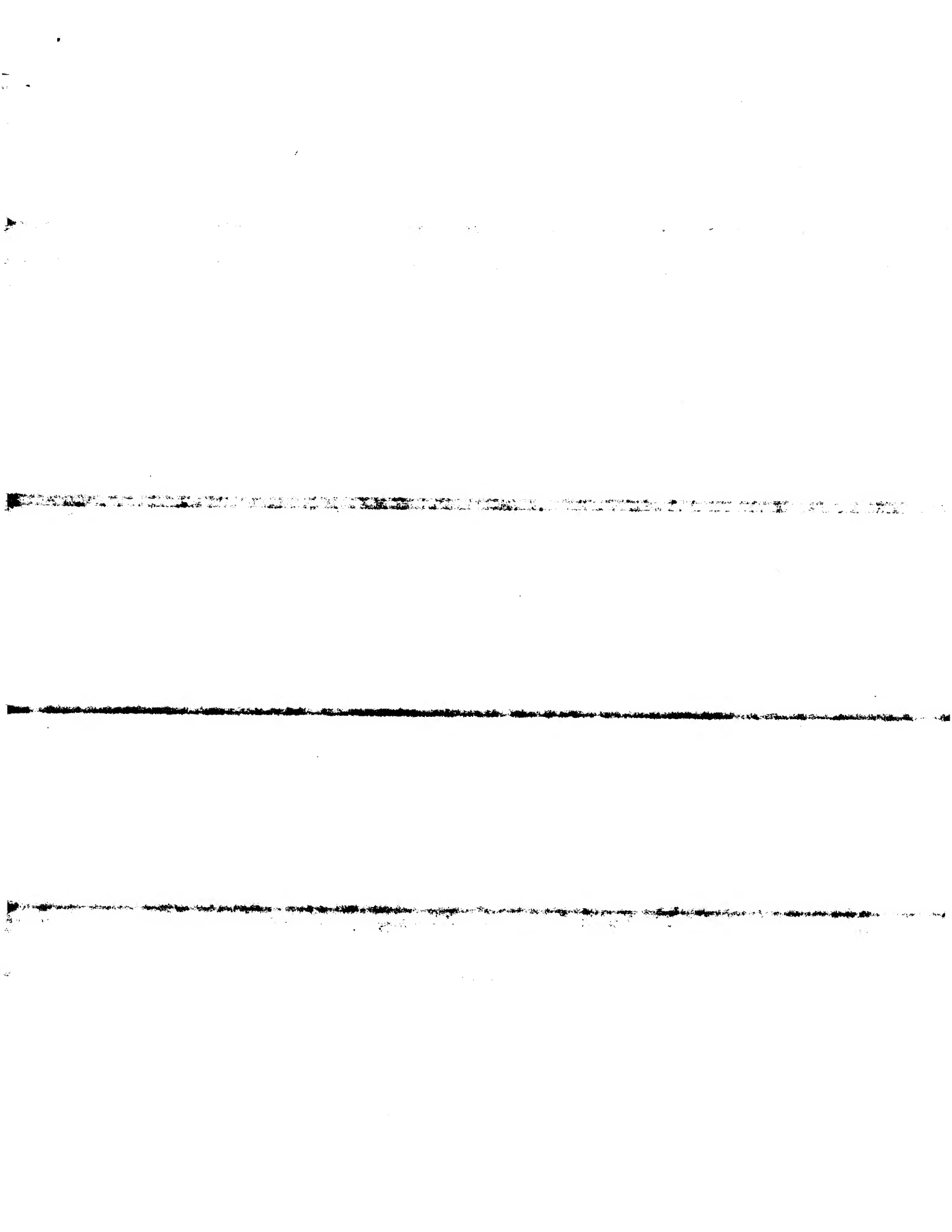
```
Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|:| || | | :|||||||| || || :|| |||||||||||:||||:|:|: :||||
Db     134 VQLAPYAKRMPHELSGGQQQRVGLARALVINPKVLLMDEPLSNLDAKLRVDLRRELREIQ 193

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             : |||||||||||||:|:|| : ||  |||||   | |:|: ||| |||:|:|  | :
Db     194 QAMNITTVYVTHDQEEALAMSDLVCVMYGGVIQQAAPPWEVYNNPANRFVASFVGANNFL 253

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQV-EEQAIHVSIRPEEFI-------KDESGDIE 292
             :   |: |   : |   : :|  | : ::: |  :||||          ||  ::
Db     254 TLD---RTGGQTSISS--LKVTLPQADAISKDRTIVAAIRPEAISIGPSIGNCDERIELP 308

Qy     293 GTIRDSVYLG--LNTDYFIETGFASK----------IQVSEESTF 325
             |||   : |  :|    : :|   :           :| ::::||
Db     309 VTIRQVSFTGREMNVAAVLSSGEEIEAITKPSPEIIALQPNQKTTF 354


RESULT   5
A72279
sugar ABC transporter, ATP-binding protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: A72279
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316
A;Accession: A72279
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-355 <ARN>
A;Cross-references: GB:AE001779; GB:AE000512; NID:g4981777; PIDN:AAD36307.1; PID:g498178
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1232
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology
F;19-210/Domain: ATP-binding cassette homology <ABC>


  Query Match             38.8%;  Score 714.5;  DB 2;  Length 355;
  Best Local Similarity   39.4%;  Pred. No. 7e-43;
  Matches  142;  Conservative   82;  Mismatches  125;  Indels   11;  Gaps    5;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |:::||   || : :|  :: ::: : :|    ||| |||||||||| |||   : ||:
Db       1 MAQVKIDGVKKYFGNVRALDGIDLVVNEGEFLVLLGPSGCGKTTLLRCIAGLEQVTGGKI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           :|:|  : |: |  ||| ||||:||::||: | ||:|: |  |||||||: ::     :|
Db      61 FFNDRDVTNLPPKDRNISMVFQSYAVWPHMKVYDNIAYPLKLKKVPKEEIEKRVKWAADL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           : |::  || | :|||||:||| :| |:   | ||||||||||:| ||: ||  ::::|
Db     121 LHISELLDRYPAQLSGGQRQRVAVARAIVHEPEVLLMDEPLSNLDALLRVKMRSELKKLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :|:||:|||||| ||| : |:|||  | :||:| | |:|| | | ||| |:|   :
Db     181 ERIGVTTIYVTHDQTEAMTMGDRIAVMNQGKLQQVGTPSEIYHHPVNIFVAGFVGSPQMN 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGD---IEGTIRD 297
             :| ||:|  :|  :|  :::||      : : : ||||    :|  :   :||
Db     241 FLEMEVRSEGNSVVLQNG-EIKIPAKTDPGAKKVILGIRPENVYLEEKPNTLKLEG---- 295

Qy     298 SVYLG--LNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNL 355
            ||    | :|  :     |: ::  :   : | : | :|   :: :|:::|  :  : :
Db     296 EVYFAEKLMSDTILHLNVGSE-KIVAKIPGDVDFRSGEKITFFLDVEKIHLFHPETGERI 354


RESULT   6
A72261
spermidine/putrescine ABC transporter, ATP-binding protein - Thermotoga maritima (strain
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: A72261
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316
A;Accession: A72261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <ARN>
A;Cross-references: GB:AE001791; GB:AE000512; NID:g4981929; PIDN:AAD36446.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1376
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
F;21-212/Domain: ATP-binding cassette homology <ABC>


  Query Match             38.0%;  Score 701;  DB 2;  Length 368;
  Best Local Similarity   42.2%;  Pred. No. 6.5e-42;
  Matches  138;  Conservative   75;  Mismatches  102;  Indels   12;  Gaps    4;

Qy       3 EIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYF 62
           |: | |  | : |  |::|::: | ||  |::|| ||||||||||:|||| :| |:
Db       5 EVSIKNVSKFFDDFQVLKNVSLDIKKGEFFSILGPSGCGKTTLLRVIAGFEGVESGDVLL 64

Qy      63 DDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQ 122
           |   | |: |:|| : ::|||||:|||||| :|:|| |  ||: : |: |: |: | |::
Db      65 DGKSILNLPPNKRPVNIIFQNYALFPHLTVFENIAFPLKLKKLSENEINQRVNELLSLIR 124

Qy     123 IAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHE 182
           : ::| : | :|||||:||| :| |||  | |||:|||||  |:|||| ::   :  :
Db     125 MEEHAQKMPSQLSGGQKQRVAIARALANEPRVLLLDEPLSALDAKLRQELLVELDNLHDR 184

Qy     183 VGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPA 242
           |||| :|||||| ||:::||::|:| :| | |:| | |:|  | | | |||||  ||:: |
Db     185 VGITFIYVTHDQAEAISVSDRVALMNEGEIVQVGTPYEVYESPVNVFAATFIGETNLMKA 244

Qy     243 NLEKRSDGAYIVFSDGYALRMPALDQVEEQA--IHVSIRPE-------EFIKDESGDI-E 292
             : :  |  |:| | |        |:  ::   : :::|||       :|  :|: ::
Db     245 EVVEVEDEYYVVESPGIGQFRCYRDKEAKKGDRLLITLRPEKIRISRKQFRSEETFNVFH 304

Qy     293 GTIRDSVYLGLNTDYFI--ETGFASKI 317
           | : : :|:|  | ||:  : |:  |:
Db     305 GVVDEEIYMGHQTKYFVRLDEGYIMKV 331


RESULT   7
H97617
ABC transporter, ATP-binding protein VCA0602 [imported] - Agrobacterium tumefaciens (
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C;Accession: H97617
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: H97617
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87897.1; PID:g15157291; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_3896
A;Map position: circular chromosome


  Query Match             36.4%;  Score 671;  DB 2;  Length 352;
  Best Local Similarity   44.1%;  Pred. No. 7.9e-40;
```

```
  Matches  146;  Conservative   54;  Mismatches  113;  Indels   18;  Gaps    4;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           || : : | || :  | |: : |: | ||   : || |||||||:|||||||    ||
Db       1 MSFLTLSNIKKSFGSVQVVHDFNMAIEKGEFVSFLGPSGCGKTTVLRMIAGFEIPTGGSI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             :      : |::|||||||| ||:||:: | :|||||     || |:  :  : |:|
Db      61 VINGKDQTTLRPNQRNIGMVFQAYALFPNMNVYENVAFGLKVAGKPKAEIDARVKEMLQL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           : :   ||| | ::|||||||| || |||  | |||:|||| |:||:|: :|: ||:||
Db     121 IHLEHLADRYPYQMSGGQQQRVALARALAPKPEVLLLDEPLSALDAKIRVSLREEIRQIQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
           ::||||::|||||||||::|||:| ||  |   ||| | ::|:||| |||:|:|  |:|
Db     181 QKLGITTIFVTHDQEEALSISDRIVVMNGGRADQIGSPFDIYNKPATRFVASFVGTLNLI 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPE--EFIKDESGD--IEGTIR 296
           | :   |     |      | | :|   : | :::|||     |  ||  | | :
Db     241 DATVVDSSSNTIRVGEQQITLPKP-VDAANGEKITLALRPEAGSLGSDAKGDVAISGLVT 299

Qy     297 DSVYLG------------LNTDYFIETGFA 314
           | :||            |: | | : | |
Db     300 SSQFLGSVIRTRLDLGGSTLSFDMFNDPGTA 330


RESULT   8
AF2840
hypothetical protein Atu2150 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AF2840
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AF2840
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL43140.1; PID:g17740615; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2150
A;Map position: circular chromosome


  Query Match             36.4%;  Score 671;  DB 2;  Length 352;
  Best Local Similarity   44.1%;  Pred. No. 7.9e-40;
  Matches  146;  Conservative   54;  Mismatches  113;  Indels   18;  Gaps    4;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           || : : | || :  | |: : |: | ||   : || |||||||:|||||||    ||
Db       1 MSFLTLSNIKKSFGSVQVVHDFNMAIEKGEFVSFLGPSGCGKTTVLRMIAGFEIPTGGSI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             :      : |::|||||||| ||:||:: | :|||||     || |:  :  : |:|
Db      61 VINGKDQTTLRPNQRNIGMVFQAYALFPNMNVYENVAFGLKVAGKPKAEIDARVKEMLQL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           : :   ||| | ::|||||||| || |||  | |||:|||| |:||:|: :|: ||:||
Db     121 IHLEHLADRYPYQMSGGQQQRVALARALAPKPEVLLLDEPLSALDAKIRVSLREEIRQIQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
           ::||||::|||||||||::|||:| ||  |   ||| | ::|:||| |||:|:|  |:|
Db     181 QKLGITTIFVTHDQEEALSISDRIVVMNGGRADQIGSPFDIYNKPATRFVASFVGTLNLI 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPE--EFIKDESGD--IEGTIR 296
           | :   |     |      | | :|   : | :::|||     |  ||  | | :
Db     241 DATVVDSSSNTIRVGEQQITLPKP-VDAANGEKITLALRPEAGSLGSDAKGDVAISGLVT 299

Qy     297 DSVYLG------------LNTDYFIETGFA 314
           | :||            |: | | : | |
Db     300 SSQFLGSVIRTRLDLGGSTLSFDMFNDPGTA 330


RESULT   9
H72555
probable transporter ATP-binding protein APE1732 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: H72555
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A;Reference number: A72450; MUID:99310339
A;Accession: H72555
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <KAW>
A;Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80733.1; PID:g5105420
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1732
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology
F;19-210/Domain: ATP-binding cassette homology <ABC>


  Query Match             36.4%;  Score 670.5;  DB 2;  Length 358;
  Best Local Similarity   39.6%;  Pred. No. 8.8e-40;
  Matches  145;  Conservative   76;  Mismatches  122;  Indels   23;  Gaps    8;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
             |: |:: :  | : :   :: ::: |  | :|||| |||||| ||:|||| : : |
Db       1 MAGIRLESVTKRFGNTVALDRVSLNISDGEIFTLLGPSGCGKTTTLRVIAGFETPDEGRV 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           |     :  ::| :||  |||||||::||: | ||:|:||  :|:|: |::::     ||
Db      61 YIGSRDVTMLKPYERNTAMVFQNYALWPHMRVFDNIAYGLKLRKLPRSEIVRRVRWAAEL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           ::|    || | :|||||||| :| |:   | |||||||||||:| ||| ||: |  :|
Db     121 LEIDHLLDRYPHQLSGGQQQRVAVARAIVTEPEVLLMDEPLSNLDAHLRLKMREEIVRLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :|:| ||||||||||::|| ::|||  | ::|:| | |:| ||| ||||||||: ::
Db     181 KRLGVTIVYVTHDQEEALSISHRVAVMNRGRVEQVGTPMEVYEKPATYFVATFIGRSTVL 240

Qy     241 PANL-EKRSDGAYIVFSDGYALRMPALDQV----EEQAIHVSIRPEEFIK---DESGD-- 290
           : |    |   |  :|  | :   |      : : : | ||||  :|   : :|:
Db     241 DGRVSEVLGSGMVRVALEG-GLSIVGTDMEGGLRQGERVKVVIRPER-VKVGHEYNGENV 298

Qy     291 IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQK----GNRIRLRINTQKLNI 346
           || :  :::||  |   :| |        :| |   | ::    |  :|  |: ::  :
Db     299 FEGKVSLAMFLGWRTQLKVEVG-------GQEITIYSDPRRAPLPGQPVRFYIDPEEAKV 351

Qy     347 FSADGS 352
           :   ||
Db     352 YRQAGS 357


RESULT  10
A82398
maltose/maltodextrin ABC transporter, ATP-binding protein VCA0946 [imported] - Vibrio
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: A82398
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
```

```
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: A82398
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <HEI>
A;Cross-references: GB:AE004421; GB:AE003853; NID:g9658375; PIDN:AAF96842.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0946
A;Map position: 2
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology


  Query Match             36.2%;  Score 668;  DB 2;  Length 373;
  Best Local Similarity   38.4%;  Pred. No. 1.4e-39;
  Matches  138;  Conservative   76;  Mismatches  135;  Indels   10;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: : : |  | | || : :|::: | :|     :| |||||:|||| |||   |  |:
Db       1 MASVTLKNVCKAYGDVLISKNVDLEIQEGEFVVFVGPSGCGKSTLLRCIAGLEDITSGDL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           |  : ::|::|||||| :||||:|||::||| : ||::|||   |  | |: :: :   |:
Db      61 YIGEQRMNDVEPSKRGVGMVFQSYALYPHLNLYDNMSFGLKLSKADKSEIKKRVDHAAEI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|::   ||:|  |||||:||| :   |   |:| |:|||||||:| ||: ||  | ::|
Db     121 LQLSHLLDRQPKALSGGQRQRVAIGRTLVSQPNVFLLDEPLSNLDAALRVQMRSEITKLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
           ::| | :|||||| ||| ::|:| |:  | : |:|:| |||| | | ||| |||  : |
Db     181 RKLGCTMIYVTHDQVEAMTMADKIVVLDAGFVSQVGKPLELYHYPQNRFVAGFIGSPKMN 240

Qy     239 IIPANLE-KRSDGAYIVFSDGYALRMP--ALDQVEEQAIHVSIRPEEFIKDESGD--IEG 293
           :  :|    |   :  |:|   :|       : : : |||  :: | ||  |||
Db     241 FMSVFIEGVEKDRVQVQLSNGTTFWIPVDGTTVTRGERMSLGIRPEHLVEAEHGDAKIEG 300

Qy     294 TIRDSVYLGLNTDYFIE-TGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           :    || | ::  | | :   :  |   |:: |: : : |   : ::| :||
Db     301 KVMIVEKLGNETQVYMNLKGSDSDVIYRQPDTL--DVETGDTLTIGIPAHRCHLFHSDG 357


RESULT  11
B98169
ATP-binding transport protein smoK (polK) AGR_L_604 [imported] - Agrobacterium tumefacie
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: B98169
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:11743194
A;Accession: B98169
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88876.1; PID:g15158643; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_604
A;Map position: linear chromosome


  Query Match             36.2%;  Score 666.5;  DB 2;  Length 346;
  Best Local Similarity   41.4%;  Pred. No. 1.6e-39;
  Matches  143;  Conservative   64;  Mismatches  115;  Indels   23;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |||::| |  | |      :  ::||: :|   |:|| |||||:||||:: | :   |||
Db       1 MSELEISNISKDYGASRALHPVSITVERGEFVTILGPSGCGKSTLLRILTGISQPSGGEI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           :|:  :|  |:| ||||:||:|||::|  |: |||  ||| |:|  ::    ||:
Db      61 RLGGKRIDQAQPEARDIAMVFQSYALFPHMSVAKNLGFGLKMKKVAKDERARRIAHALEI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :    || | :|||||||| || |: : ||:|| |||||||:|||| :|  : |:
Db     121 CNLTGLVDRMPRQLSGGQQQRVALARAIVMQPSLLLFDEPLSNLDAKLRDTLRHELTELH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :| |::|||||| ||||:||:| ||  | : :|| | |||  | : | | |:|:||::
Db     181 RRIGATSLYVTHDQAEAMAMSDRIVVMNAGRVVEIGTPLELYRAPKHAFTAGFLGQTNLL 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMP-----ALDQVEEQAIHVSIRPE--EFIKDESGDIEG 293
           |             | ::|   ::|       |||      : :| |||      |::||  |
Db     241 P------------VSAEGQQAQLPWGQSVTLDQAAAGNVQISARPENIHICADQAGD--G 286

Qy     294 TIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLR 338
           |:    ::| |  | :| |   :|:|| :|  |  :  | |: |:
Db     287 TVSAVSFMGANALYTVEIG-GRQIRVS-QSGAETLIDAGQRVALQ 329


RESULT  12
AE3118
hypothetical protein Atu4569 [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AE3118
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AE3118
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45363.1; PID:g17743058; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4569
A;Map position: linear chromosome


  Query Match             36.2%;  Score 666.5;  DB 2;  Length 346;
  Best Local Similarity   41.4%;  Pred. No. 1.6e-39;
  Matches  143;  Conservative   64;  Mismatches  115;  Indels   23;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |||::| |  | |      :  ::||: :|   |:|| |||||:||||:: | :   |||
Db       1 MSELEISNISKDYGASRALHPVSITVERGEFVTILGPSGCGKSTLLRILTGISQPSGGEI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           :|:  :|  |:| ||||:||:|||::|  |: |||  ||| |:|  ::    ||:
Db      61 RLGGKRIDQAQPEARDIAMVFQSYALFPHMSVAKNLGFGLKMKKVAKDERARRIAHALEI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :    || | :|||||||| || |: : ||:|| |||||||:|||| :|  : |:
Db     121 CNLTGLVDRMPRQLSGGQQQRVALARAIVMQPSLLLFDEPLSNLDAKLRDTLRHELTELH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :| |::|||||| ||||:||:| ||  | : :|| | |||  | : | | |:|:||::
Db     181 RRIGATSLYVTHDQAEAMAMSDRIVVMNAGRVVEIGTPLELYRAPKHAFTAGFLGQTNLL 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMP-----ALDQVEEQAIHVSIRPE--EFIKDESGDIEG 293
           |             | ::|   ::|       |||      : :| |||      |::||  |
Db     241 P------------VSAEGQQAQLPWGQSVTLDQAAAGNVQISARPENIHICADQAGD--G 286

Qy     294 TIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLR 338
           |:    ::| |  | :| |   :|:|| :|  |  :  | |: |:
Db     287 TVSAVSFMGANALYTVEIG-GRQIRVS-QSGAETLIDAGQRVALQ 329
```

```
RESULT  13
AH3583
glucose ABC transporter ATPase BMEII0593 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AH3583
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AH3583
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53835.1; PID:g17984770; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0593
A;Map position: II


  Query Match             35.9%;  Score 662;  DB 2;  Length 373;
  Best Local Similarity   40.2%;  Pred. No. 3.7e-39;
  Matches  150;  Conservative   62;  Mismatches  141;  Indels   20;  Gaps    7;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           || :|| :  | |  | |:::::|| | :|    |:| |||||:|||  |||   |  |:
Db       3 MSFLKITDLYKSYGSVSVLKDINIEIDEGGFLVLVGPSGCGKSTLLNTIAGLEPITSGDI 62

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             :  ::|: ||:|:| ||||:||::|::||  |:|||:  :|||| |  :   :  ::
Db      63 AINGKSVSNLHPSQRDIAMVFQSYALYPNMTVAGNIAFGMEIRKVPKPERDKAIQQVADM 122

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :||    |||| :|||||::|| :  ||  || | | ||||||:||||:|||  |: :
Db     123 LQIGHLLDRKPSQLSGGQRRRVAMGRALVRNPQVFLFDEPLSNLDAKLRVDMRTEIKRLH 182

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
             :  | ||||||| ||| :: :|||:||||:|| | | |:|: ||| ||| |:|    |
Db     183 SRMKTTIVYVTHDQIEAMTLATKIAVLKDGVLQQFGTPAEIYNNPANMFVADFMGSPAMN 242

Qy     239 IIPANLEKRSDGAYIVFSD---GYALRMP------ALDQVEEQAIHVSIRPEEFIKDESG 289
           ::|  :|| |   | :  |   |  |::|       | |   : :   ||||    :
Db     243 LLPVKVEK-SGSDYAITLDQGEGEVLKLPLKAVPQGLAQYVGKEVVFGIRPEALTDPDGA 301

Qy     290 D------IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQK 343
           |      :||     |     :| |  |    |  |       :  :: |   ||  |  |
Db     302 DRNAASIVEGDCLIDVVEPAGSDTFAVTRLGGK-HVVARLRADARIRAGQSARLAFNLDK 360

Qy     344 LNIFSADGSQNLI 356
            ::|    ||  |
Db     361 -SVFFDSASQTRI 372


RESULT  14
AD3587
SN-glycerol-3-phosphate transport ATP-binding protein ugpC BMEII0621 [imported] - Brucel
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AD3587
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AD3587
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53863.1; PID:g17984801; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0621
A;Map position: II


  Query Match             35.9%;  Score 661.5;  DB 2;  Length 351;
  Best Local Similarity   40.2%;  Pred. No. 3.7e-39;
  Matches  148;  Conservative   70;  Mismatches  111;  Indels   39;  Gaps    8;

Qy       1 MSEIKIINAKKIY-HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
           ||:| : | :| |  :: ||: ::: |  |    |:| |||||:|||||||  ||  |
Db       1 MSKIVLDNVRKSYGGNIEVIKGVSLEIADGEFVVLVGPSGCGKSTLLRMIAGLESITSGT 60

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
                :  :||:||::|:| |||||||::||:|||:|:|:||  :| ||||: ::  |  :
Db      61 ISIGERVVNNVEPAERDIAMVFQNYALYPHMTVRENLAYGLKNRKTPKEEIERRIAKAAK 120

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :::| |: :||| :|||||:||| :  |:   |:  | |||||||:||||: ||  |: :
Db     121 VLEIEQFLERKPRQLSGGQRQRVAMGRAIVREPAAFLFDEPLSNLDAKLRVQMRVEIKRL 180

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNI 239
           |  :| |:|||||| ||| ::|:: |:  | |:|:| | ||| |||: |||||||  ::
Db     181 QRSLGTTSVYVTHDQMEAMTMADRLVVLNAGHIEQVGTPIELYEKPASIFVATFIGSPSM 240

Qy     240 IPANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSV 299
             || :  : |   :  | |: :|:            :|||:    | || :
Db     241 ---NLLQSPESA--AWQPGRAITLPS------GGYTFGVRPEDIRILEEGDQDAD----- 284

Qy     300 YLGLNTDYFIE----TGFASKI-----------QVSEESTFEEDLQKGNRIRLRINTQK 343
             | |    ||      |  | |                :|:  ||   |      :|:  :
Db     285 --GFNAQVRIEAVELVGAESYIHAALSDGKPLIFRVAGRSTHNID----EMVRVGASATD 338

Qy     344 LNIFSADG 351
           ::|| |||
Db     339 VHIFGADG 346


RESULT  15
T35802
probable polyamine ABC-transporter ATP-binding protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 17-Mar-2000
C;Accession: T35802
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z21556
A;Accession: T35802
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-397 <MUR>
A;Cross-references: EMBL:AL031225; PIDN:CAA20221.1; GSPDB:GN00070; SCOEDB:SC8B7.10c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC8B7.10c
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog


  Query Match             35.9%;  Score 661.5;  DB 2;  Length 397;
  Best Local Similarity   38.3%;  Pred. No. 4.4e-39;
  Matches  142;  Conservative   76;  Mismatches  134;  Indels   19;  Gaps    6;

Qy       3 EIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYF 62
           ::::    | |     :  |::|:|:|| | |||||||||||| ||||||     ||  :
Db      18 DVRLAGISKTYGSFTAVHPLDLTVPQGSFFALLGASGCGKTTTLRMIAGLEEPSGGTVHL 77

Qy      63 DDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQ 122
             |  :  : | || :  |||:||:|||| : :|||||| :: :  : : :|  : |:|:|
Db      78 GDQDVTALPPYKRPVNTVFQSYALFPHLDIYENVAFGLRRRGI--KSVKKQVGEMLDLVQ 135

Qy     123 IAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHE 182
           : : | :|| :|||||||||| :| ||  :| |||:||||  |: |||  |:  :: || |
```

```
Db     136 LGEQARKKPHQLSGGQQQRVAVARALINHPKVLLLDEPLGALDLKLRRQMQLELKRIQTE 195

Qy     183 VGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPA 242
           |||| |:|||||||||| ::||:|||  | ::|:| | :||  |   ||| |:| :|:| |
Db     196 VGITFVHVTHDQEEAMTMADQVAVMNAGRVEQLGAPADLYENPRTTFVANFLGTSNLIEA 255

Qy     243 NLEKRSDGAYIVFSDGYALRMP----ALDQVEEQAIHVSIRPEEFI---KDESGD----- 290
            :: ||    :| : |  | :|    :        : | :|||:      |::|
Db     256 EVDTRSGDDVVVKAAGDKLVLPGARCSAPAKTGGKVLVGVRPEKISLTHADDAGSIPEGR 315

Qy     291 --IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEED-LQKGNRIRLRINTQKLNIF 347
             | | |  : ::|::| | ::     : :|  ::   :  |  |  : |  |    : |
Db     316 NRITGRIGSTSFIGVSTQYVVDCAACPEFEVYAQNIDRDPRLTPGAEVVLHWN--PAHTF 373

Qy     348 SADGSQNLIKG 358
           |  |::: |
Db     374 GLDADQSVLAG 384


Search completed: July 30, 2002, 15:48:10
Job time: 144 sec
```



GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:48:12 ; Search time 13.45 Seconds
(without alignments)
1044.996 Million cell updates/sec

Title: US-09-769-787-162
Perfect score: 1843
Sequence: 1 MSEIKIINAKKIYHDVPVIE.....LNIFSADGSQNLIKGVNHGT 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	651.5	35.3	381	1	POTA_HAEIN	P45171 haemophilus
2	644	34.9	378	1	POTA_ECOLI	P23858 escherichia
3	638.5	34.6	348	1	AFUC_ECOLI	P37009 escherichia
4	638	34.6	363	1	LACK_AGRRD	Q01937 agrobacteri
5	635	34.5	362	1	AGLK_RHIME	Q9z3r9 rhizobium m
6	629	34.1	369	1	MALK_SALTY	P19566 salmonella
7	629	34.1	378	1	POTA_SALTY	P40790 salmonella
8	628	34.1	371	1	MALK_ECOLI	P02914 escherichia
9	627.5	34.0	328	1	AFUC_HAEIN	P44531 haemophilus
10	627	34.0	377	1	POTG_ECOLI	P31134 escherichia
11	611.5	33.2	365	1	MSMX_BACSU	P94360 bacillus su
12	602.5	32.7	322	1	YCJV_ECOLI	P77481 escherichia
13	600	32.6	355	1	CYSA_SYNY3	P74548 synechocyst
14	599	32.5	332	1	SMOK_RHOSH	P54933 rhodobacter
15	592	32.1	371	1	Y4OS_RHISN	P55604 rhizobium s
16	590	32.0	377	1	MSMK_STRMU	Q00752 streptococc
17	586.5	31.8	344	1	CYSA_SYNP7	P14788 synechococc
18	584	31.7	348	1	AFUC_ACTPL	Q57293 actinobacil
19	576.5	31.3	337	1	YDCT_ECOLI	P77795 escherichia
20	564	30.6	294	1	MALK_ENTAE	P18813 enterobacte
21	550	29.8	348	1	CYSA_MESVI	Q9mun1 mesostigma
22	544	29.5	356	1	UGPC_ECOLI	P10907 escherichia
23	532	28.9	236	1	CYSA_CHLVU	P56344 chlorella v
24	531	28.8	365	1	CYSA_ECOLI	P16676 escherichia
25	518.5	28.1	339	1	Y4FO_RHISN	P55453 rhizobium s
26	518	28.1	365	1	CYSA_SALTY	P40860 salmonella
27	513.5	27.9	586	1	Y187_MYCPN	P75264 mycoplasma
28	494	26.8	585	1	Y1B7_MYCGE	P47433 mycoplasma
29	489.5	26.6	356	1	HITC_HAEIN	P44513 haemophilus
30	477	25.9	345	1	SFUC_SERMA	P21410 serratia ma
31	472	25.6	370	1	CYSA_MARPO	P10091 marchantia
32	446.5	24.2	559	1	POTA_MYCGE	P47288 mycoplasma
33	440	23.9	418	1	OPAA_BACSU	P46920 bacillus su
34	431	23.4	380	1	OPCA_BACSU	O34992 bacillus su
35	427	23.2	381	1	OPBA_BACSU	Q45460 bacillus su
36	423	23.0	400	1	PROV_ECOLI	P14175 escherichia
37	421	22.8	400	1	PROV_SALTY	P17328 salmonella
38	420	22.8	560	1	POTA_MYCPN	P75059 mycoplasma
39	412.5	22.4	297	1	YD67_METJA	Q58762 methanococc
40	394	21.4	242	1	GLNQ_BACST	P27675 bacillus st
41	389	21.1	242	1	GLUA_CORGL	P48243 corynebacte
42	375	20.3	263	1	BZTD_RHOCA	Q52666 rhodobacter
43	372	20.2	308	1	YEHX_ECOLI	P33360 escherichia
44	368	20.0	240	1	GLNQ_ECOLI	P10346 escherichia
45	367.5	19.9	267	1	Y412_METJA	Q57855 methanococc

ALIGNMENTS

```
RESULT   1
POTA_HAEIN
ID   POTA_HAEIN     STANDARD;      PRT;   381 AA.
AC   P45171;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Spermidine/putrescine transport ATP-binding protein potA.
GN   POTA OR HI1347.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY
CC       COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U32813; AAC22991.1; -.
DR   TIGR; HI1347; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; ATP-binding; Inner membrane; Complete proteome.
FT   NP_BIND      52      59       ATP (POTENTIAL).
SQ   SEQUENCE   381 AA;  43429 MW;  3E61DD3D062EBECA CRC64;
```

```
  Query Match             35.3%;  Score 651.5;  DB 1;  Length 381;
  Best Local Similarity   40.6%;  Pred. No. 9.2e-40;
  Matches  139;  Conservative   70;  Mismatches  106;  Indels   27;  Gaps    7;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           |:: : || |    :| : |:||  |   |:|| |||||||:||::||   :: |    |
Db      20 IELRSIKKSYGSNTIINDFNLTINNGEFVTILGPSGCGKTTVLRLLAGLEELDSGSIILD 79

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             | |:   ||:|  |||:||:|||:|: :|||||  :||| ||:  :  : | ::|:
Db      80 GEDITNVPAEKRHINTVFQSYALFPHMTIFENVAFGLRMQKVPNEEIKPRVLEALRMVQL 139

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           : ||||| :|||||||||: :| |:   | |||:|| || |: |||  |:| :: :| ::
Db     140 EEMADRKPTQLSGGQQQRIAIARAVVNKPKVLLLDESLSALDYKLRKQMQQELKMLQRQL 199

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           ||| ::|||||||||: :||:| ::: | | | | |:|:|  |||  |||  |||  |:  |
Db     200 GITFIFVTHDQEEAITMSDRIVLLRKGKIAQDGSPREIYEDPANLFVARFIGEINVFEAT 259

Qy     244 -LEKRSDGAYI---------VFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDE------ 287
            :|::|:   :          :::|     ||     ::| : | :|||: : :|
Db     260 VIERKSEQVVLANVEGRICDIYTD-----MPV---EKDQKLQVLLRPEDIVIEELDENEH 311

Qy     288 SGDIEGTIRDSVYLGLNTDYFIETGF-ASKIQVSEESTFEED 328
           |  | `| | |  | |:  :  :|      :: |||   | ||
Db     312 SKAIIGHIIDRTYKGMTLESTVEFDHNGMRVLVSE--FFNED 351


RESULT   2
POTA_ECOLI
ID   POTA_ECOLI     STANDARD;      PRT;   378 AA.
AC   P23858;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Spermidine/putrescine transport ATP-binding protein potA.
GN   POTA OR B1126 OR Z1831 OR ECS1571.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92041956; PubMed=1939142;
RA   Furuchi T., Kashiwagi K., Kobayashi H., Igarashi K.;
RT   "Characteristics of the gene for a spermidine and putrescine
RT   transport system that maps at 15 min on the Escherichia coli
RT   chromosome.";
RL   J. Biol. Chem. 266:20928-20933(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY
CC       COUPLING TO THE TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M64519; AAC37038.1; -.
DR   EMBL; AE000212; AAC74210.1; -.
DR   EMBL; D90747; BAA35946.1; -.
DR   EMBL; D90748; BAA35948.1; -.
DR   EMBL; AE005326; AAG55930.1; -.
DR   EMBL; AP002555; BAB34994.1; -.
DR   PIR; A40840; A40840.
DR   EcoGene; EG10749; potA.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; ATP-binding; Inner membrane; Complete proteome.
FT   NP_BIND      50     57       ATP (BY SIMILARITY).
SQ   SEQUENCE   378 AA;  43028 MW;  14DCA99329A344F3 CRC64;


  Query Match             34.9%;  Score 644;  DB 1;  Length 378;
  Best Local Similarity   39.7%;  Pred. No. 3.1e-39;
  Matches  137;  Conservative   65;  Mismatches  109;  Indels   34;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           :::   :| :    ||  |::||  |   |||| |||||||:||:|||  ::: |    |
Db      18 VQLAGIRKCFDGKEVIPQLDLTINNGEFLTLLGPSGCGKTTVLRLIAGLETVDSGRIMLD 77

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
           :  | ::    | :  |||:||:|||:|| :||||||  :| |  |:  :  : | ::|:
Db      78 NEDITHVPAENRYVNTVFQSYALFPHMTVFENVAFGLRMQKTPAAEITPRVMEALRMVQL 137

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
             :| ||| :||||||||| :| |:   | :||:|| || |: |||  |:  :: :| ::
Db     138 ETFAQRKPHQLSGGQQQRVAIARAVVNKPRLLLLDESLSALDYKLRKQMQNELKALQRKL 197
```

```
Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNI---- 239
           ||| |:|||||||||: :||:| ||:|| |:| | |:|:| :| | ||| |||  |:
Db     198 GITFVFVTHDQEEALTMSDRIVVMRDGRIEQDGTPREIYEEPKNLFVAGFIGEINMFNAT 257

Qy     240 ---------IPANLEKRSDGAYIVFSDGYALRMPALDQVEE-QAIHVSIRPEEFIKDESG 289
                   : ||:| |    |: |:           ||  | :|| :|||:   :|
Db     258 VIERLDEQRVRANVEGRECNIYVNFA-----------VEPGQKLHVLLRPEDLRVEEIN 305

Qy     290 D------IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEED 328
           |      : | :|:  | |:  :  :|      : |||   | ||
Db     306 DDNHAEGLIGYVRERNYKGMTLESVVELENGKMVMVSE--FFNED 348


RESULT   3
AFUC_ECOLI
ID   AFUC_ECOLI     STANDARD;      PRT;   348 AA.
AC   P37009; P77157;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ferric transport ATP-binding protein afuC.
GN   AFUC OR B0262.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / W3110;
RA   Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA   Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA   Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT   "Systematic sequencing of the Escherichia coli genome: analysis of the
RT   4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA   Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA   Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA   Davis R.W.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-220 FROM N.A.
RX   MEDLINE=94156832; PubMed=8113168;
RA   Volkert M.R., Loewen P.C., Switala J., Crowley D., Conley M.;
RT   "The delta (argF-lacZ)205(U169) deletion greatly enhances resistance
RT   to hydrogen peroxide in stationary-phase Escherichia coli.";
RL   J. Bacteriol. 176:1297-1302(1994).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR IRON. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- MISCELLANEOUS: IN E.COLI, THE AFU SYSTEM SEEMS TO HAVE BEEN
CC       DELETED BY AN INSERTION SEQUENCE. AFUA IS TOTALLY LOST, AFUB IS
CC       PARTIALLY PRESENT AND AFUC IS TOTALLY CONSERVED.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   -!- CAUTION: REF.4 DIFFERS FROM THAT SHOWN FROM POSITION 147 TO 220
CC       DUE TO A FRAMESHIFT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000134; AAC73365.1; ALT_INIT.
DR   EMBL; D83536; BAA77930.1; -.
DR   EMBL; U70214; AAB08683.1; ALT_INIT.
DR   EMBL; L20943; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG12340; afuC.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Iron transport; Transport; ATP-binding; Complete proteome.
FT   NP_BIND      39     46       ATP (POTENTIAL).
FT   CONFLICT    147    148       LA -> WP (IN REF. 2).
SQ   SEQUENCE   348 AA;  39059 MW;  D4AD33754217DA1A CRC64;


  Query Match             34.6%;  Score 638.5;  DB 1;  Length 348;
  Best Local Similarity   39.7%;  Pred. No. 7e-39;
  Matches  136;  Conservative   64;  Mismatches  116;  Indels   27;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
             ::: |  | :    ||:|:|:|||:| : |||| |||||||:||::||      |: : |
Db       7 VELRNVTKRFGSNTVIDNINLTIPQGQMVTLLGPSGCGKTTILRLVAGLEKPSEGQIFID 66

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
              : :    :|:| ||||:||:|||::: :|| :||    ||: ||  :  : | :: :
Db      67 GEDVTHRSIQQRDICMVFQSYALFPHMSLGENVGYGLKMLGVPRAELKARVKEALAMVDL 126

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
             : ||  |::|||||||| || || : | ||| |||||||:| ||  ||  |||:| :
Db     127 EGFEDRFVDQISGGQQQRVALARALILKPKVLLFDEPLSNLDANLRRSMRDKIRELQKQF 186

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
               ||::|||||| || |:|| : ||  | | ||| |::|| :||: |:|:|:|  |: ||
Db     187 DITSLYVTHDQSEAFAVSDTVLVMNKGHIMQIGSPQDLYRQPASRFMASFMGDANLFPAT 246

Qy     244 LEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEG--TIRDSVYL 301
               |||   ::  || |  |     : : : | :|||     : |:      ||   |:
Db     247 F---SDGYVDIY--GYHLPRPLHFGTQGEGM-VGVRPEAITLSDRGEESQRCVIRHVAYM 300

Qy     302 GLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKL 344
           |   :  :|                    | | |::|  :|
Db     301 GPQYEVTVE------------------WHGQEILLQVNATRL 324


RESULT   4
LACK_AGRRD
ID   LACK_AGRRD     STANDARD;      PRT;   363 AA.
AC   Q01937;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lactose transport ATP-binding protein lacK.
GN   LACK.
OS   Agrobacterium radiobacter.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=358;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR50;
RX   MEDLINE=92334152; PubMed=1630315;
RA   Williams S.G., Greenwood J.A., Jones C.W.;
RT   "Molecular analysis of the lac operon encoding the binding-protein-
RT   dependent lactose transport system and beta-galactosidase in
RT   Agrobacterium radiobacter.";
RL   Mol. Microbiol. 6:1755-1768(1992).
```

```
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR LACTOSE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66596; -; NOT_ANNOTATED_CDS.
DR   PIR; S34734; S34734.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Sugar transport; Transport; ATP-binding; Inner membrane.
FT   NP_BIND      36     43       ATP (BY SIMILARITY).
SQ   SEQUENCE   363 AA;  39324 MW;  1B1F8CC29453B8B2 CRC64;


  Query Match             34.6%;  Score 638;  DB 1;  Length 363;
  Best Local Similarity   38.3%;  Pred. No. 8e-39;
  Matches  138;  Conservative   69;  Mismatches  143;  Indels   10;  Gaps    4;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |:|::: : :| |  : ||: :|: :  |     :| |||||:||||||||   |  ||
Db       1 MAEVRLTDIRKSYGSLEVIKGVNLEVSSGEFVVFVGPSGCGKSTLLRMIAGLEDISSGEL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             | :|:::|||| | |||| ||::||:|||:|: | |    : |:|: :: |   ::
Db      61 TIGGTVMNDVDPSKRGIAMVFQTYALYPHMTVRENMGFALRFAGMAKDEIERRVNAAAKI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :::    |||| |||||:||| :  |:   | | | |||||||:|:||: ||  |  :
Db     121 LELDALMDRKPKALSGGQRQRVAIGRAIVRQPDVFLFDEPLSNLDAELRVHMRVEIARLH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
            |:   | ||||||| ||| ::|:| ||: |:::|:| |  ||  | | ||| |||  | |
Db     181 KELNATIVYVTHDQVEAMTLADKIVVMRGGIVEQVGAPLALYDDPDNMFVAGFIGSPRMN 240

Qy     239 IIPANLEKRSDGAYIVFS-----DGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEG 293
            :|| :  :::|  :  :     |       |       |: | :||| |:   ||| :
Db     241 FLPAVVIGQAEGGQVTVALKARPDTQLTVACATPPQGGDAVTVGVRPEHFLPAGSGDTQL 300

Qy     294 TIRDSV--YLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           |    |  :|| || |        :  : |:       : |: | : |: :   :| | |
Db     301 TAHVDVVEHLG-NTSYVYAHTVPGEQIIIEQEERRHGGRYGDEIAVGISAKTSFLFDASG 359


RESULT   5
AGLK_RHIME
ID   AGLK_RHIME     STANDARD;      PRT;   362 AA.
AC   Q9Z3R9;
DT   30-MAY-2000 (Rel. 39, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Alpha-glucoside transport ATP-binding protein aglK.
GN   AGLK OR R00699 OR SMC03065.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99328961; PubMed=10400573;
RA   Willis L.B., Walker G.C.;
RT   "A novel Sinorhizobium meliloti operon encodes an alpha-glucosidase
RT   and a periplasmic-binding-protein-dependent transport system for
RT   alpha-glucosides.";
RL   J. Bacteriol. 181:4176-4184(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR ALPHA-GLUCOSIDES SUCH AS SUCROSE, MALTOSE AND TREHALOSE.
CC       PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF045609; AAD12046.1; -.
DR   EMBL; AL591784; CAC45271.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Sugar transport; Transport; ATP-binding; Inner membrane;
KW   Complete proteome.
FT   NP_BIND      36     43       ATP (POTENTIAL).
FT   CONFLICT    173    185       RIEIAKLSERMSD -> AIENRQAQRAMSN (IN REF.
FT                                1).
FT   CONFLICT    208    216       LSAGHIEQV -> FSRRSTFEQF (IN REF. 1).
FT   CONFLICT    220    220       L -> F (IN REF. 1).
FT   CONFLICT    229    229       L -> F (IN REF. 1).
FT   CONFLICT    239    239       A -> P (IN REF. 1).
SQ   SEQUENCE   362 AA;  39463 MW;  66B76CF47FE95A39 CRC64;


  Query Match             34.5%;  Score 635;  DB 1;  Length 362;
  Best Local Similarity   38.5%;  Pred. No. 1.3e-38;
  Matches  138;  Conservative   70;  Mismatches  142;  Indels    8;  Gaps    5;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: : : : :| |  | ||  ::: | :|     :| |||||:|||||||   | ||:
Db       1 MTGLLLKDIRKSYGAVDVIHGIDLDIKEGEFVVFVGPSGCGKSTLLRMIAGLEEITGGDM 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           : |  ::|:: |||| | ||||:||::||:|| ||:|||:   :  |||: ::     ::
Db      61 FIDGERVNDVPPSKRGIAMVFQSYALYPHMTVYDNMAFGMRIARESKEEIDRRVRGAADM 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|:  | || |  |||||:||| :  |:  || | | |||||||:| ||:  |  | ::
Db     121 LQLTPYLDRLPKALSGGQRQRVAIGRAICRNPKVFLFDEPLSNLDAALRVATRIEIAKLS 180

Qy     181 HEVGITT-VYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RT 237
             :  || :|||||| ||| ::|:| |:  | |:|:| | ||| :||| ||| |||
Db     181 ERMSDTTMIYVTHDQVEAMTLADRIVVLSAGHIEQVGAPLELYERPANLFVARFIGSPAM 240

Qy     238 NIIPANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVS--IRPEEFIKDESGD--IEG 293
           |:||| :        :  : | :: :       |     | :|||:    |: |   ||
Db     241 NVIPATITATGQQTAVSLAGGKSVTLDVPTNASENGKTASFGVRPEDLRVTEADDFLFEG 300
```

```
Qy     294 TIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           |:     ||  |  :|| |      :  :      : :|:::|   :  ||::|  :|
Db     301 TVSIVEALGEVTLLYIE-GLVENEPIIAKMPGIARVGRGDKVRFTADKAKLHLFDTNG 357


RESULT   6
MALK_SALTY
ID   MALK_SALTY     STANDARD;      PRT;   369 AA.
AC   P19566; O06952;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Maltose/maltodextrin transport ATP-binding protein malK.
GN   MALK OR STM4230.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89384443; PubMed=2674653;
RA   Dahl M.K., Francoz E., Saurin W., Boos W., Manson M.D., Hofnung M.;
RT   "Comparison of sequences from the malB regions of Salmonella
RT   typhimurium and Enterobacter aerogenes with Escherichia coli K12: a
RT   potential new regulatory site in the interoperonic region.";
RL   Mol. Gen. Genet. 218:199-207(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=92110387; PubMed=1730061;
RA   Schneider E., Francoz E., Dassa E.;
RT   "Completion of the nucleotide sequence of the 'maltose B' region in
RT   Salmonella typhimurium: the high conservation of the malM gene
RT   suggests a selected physiological role for its product.";
RL   Biochim. Biophys. Acta 1129:223-227(1992).
RN   [3]
RP   REVISIONS TO 99-102.
RA   Schneider E.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: THIS IS ONE OF THE FIVE PROTEINS ESSENTIAL TO THE ACTIVE
CC       BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEM FOR MALTOSE AND
CC       MALTODEXTRIN.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X54292; CAA38188.1; -.
DR   EMBL; AE008897; AAL23054.1; -.
DR   PIR; S05329; S05329.
DR   PIR; S20602; S20602.
DR   StyGene; SG10215; malK.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; Sugar transport; Inner membrane; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      36     43       ATP (BY SIMILARITY).
FT   CONFLICT    141    141       R -> A (IN REF. 1 AND 2).
SQ   SEQUENCE   369 AA;  40799 MW;  98CA374498C88FCF CRC64;


  Query Match             34.1%;  Score 629;  DB 1;  Length 369;
  Best Local Similarity   35.7%;  Pred. No. 3.6e-38;
  Matches  132;  Conservative   81;  Mismatches  127;  Indels   30;  Gaps    7;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: ::: |  | : || | :::|: |  |     :| |||||:||||||||  :|  |:
Db       1 MASVQLRNVTKAWGDVVVSKDINLDIHDGEFVVFVGPSGCGKSTLLRMIAGLETITSGDL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             :  :|::|:: |::| :|||||:||::|||:| :|::|||      || : |: |:  |:
Db      61 FIGETRMNDIPPAERGVGMVFQSYALYPHLSVAENMSFGLKLAGAKKEVMNQRVNQVAEV 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|:|   :|||  |||||:||| :   |   | | |:|||||||:| ||: ||  |  :
Db     121 LQLAHLLERKPKALSGGQRQRVAIGRTLVAEPRVFLLDEPLSNLDAALRVQMRIEISRLH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
             :| | :|||||| ||| ::|:| |:  | : |:|:| |||| ||: ||| |||  : |
Db     181 KRLGRTMIYVTHDQVEAMTLADKIVVLDAGRVAQVGKPLELYHYPADRFVAGFIGSPKMN 240

Qy     239 IIPANLEKRSDGAYIVFSDGYALRMPALDQ----VEEQAIHV------SIRPEEFIKDES 288
           :|  :   :        :   : :|   |    || : : |       ||||  :  :
Db     241 FLPVKVTATA-------IEQVQVELPNRQQIWLPVESRGVQVGANMSLGIRPEHLLPSDI 293

Qy     289 GDI--EGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEED----LQKGNRIRLRINTQ 342
           |:  || ::    ||  |   |:        : :   : ::    :::|    : :  :
Db     294 ADVTLEGEVQVVEQLGHETQIHIQIP-----AIRQNLVYRQNDVVLVEEGATFAIGLPPE 348

Qy     343 KLNIFSADGS 352
           : ::|  |||
Db     349 RCHLFREDGS 358


RESULT   7
POTA_SALTY
ID   POTA_SALTY     STANDARD;      PRT;   378 AA.
AC   P40790;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Spermidine/putrescine transport ATP-binding protein potA.
GN   POTA OR STM1226.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [2]
RP   SEQUENCE OF 1-63 FROM N.A.
RX   MEDLINE=91258340; PubMed=1904438;
RA   Miller C.G., Miller J.L., Bagga D.A.;
```

```
RT   "Cloning and nucleotide sequence of the anaerobically regulated pepT
RT   gene of Salmonella typhimurium.";
RL   J. Bacteriol. 173:3554-3558(1991).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY
CC       COUPLING TO THE TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC       frameshift in position 55.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008753; AAL20155.1; -.
DR   EMBL; M62725; -; NOT_ANNOTATED_CDS.
DR   StyGene; SG10489; potA.
DR   InterPro; IPR003439; ABC_transportr.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; ATP-binding; Inner membrane; Complete proteome.
FT   NP_BIND      50      57       ATP (BY SIMILARITY).
SQ   SEQUENCE   378 AA;  42840 MW;  470FEC6B00F17216 CRC64;


  Query Match             34.1%;  Score 629;  DB 1;  Length 378;
  Best Local Similarity   42.3%;  Pred. No. 3.7e-38;
  Matches  138;  Conservative   61;  Mismatches  117;  Indels   10;  Gaps    5;

Qy      11 KIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNM 70
           | :    ||  |::||  |   |||| |||||||:||:|||  ::: |    |:  | ::
Db      25 KSFDGKEVISQLDLTINNGEFLTLLGPSGCGKTTVLRLIAGLETVDAGHIMLDNQDITHV 84

Qy      71 EPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRK 130
             | :  |||:||:|||:|| :||||||  :| |  |:  :    | ::|: ::| ||
Db      85 PAENRYVNTVFQSYALFPHMTVFENVAFGLRMQKTPAAEIAPRVTDALRMVQLEEFAQRK 144

Qy     131 PDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITTVYV 190
           | :|||||||| :| |:   | :||:|| || |: |||  |:  :: :| ::||| |:|
Db     145 PHQLSGGQQQRVAIARAVVNKPRLLLLDESLSALDYKLRKQMQNELKALQRKLGITFVFV 204

Qy     191 THDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEKRSDG 250
           |||||||: :||:| ||::|||:| | |:|:| :| | ||| |||  |   | : :| |
Db     205 THDQEEALTMSDRIVVMRNGVIEQDGTPREIYEEPKNLFVAGFIGEINRFDATVIERLDE 264

Qy     251 AYIVFS-DGYALRMPALDQVEE-QAIHVSIRPEEFIKDESGD---IEGTI---RDSVYLG 302
             : | :|    :     ||  | ::| :|||:   :|  |   ||| |   |:  | |
Db     265 QRVRASVEGRECNIYVNFAVEPGQKLNVLLRPEDLRVEEINDDNHIEGLIGYVRERNYKG 324

Qy     303 LNTDYFIETGFASKIQVSEESTFEED 328
           :  :  :|      : |||   | ||
Db     325 MTLESVVELENGKMVMVSE--FFNED 348


RESULT   8
MALK_ECOLI
ID   MALK_ECOLI     STANDARD;      PRT;   371 AA.
AC   P02914; Q47348;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Maltose/maltodextrin transport ATP-binding protein malK.
GN   MALK OR B4035 OR Z5633 OR ECS5018.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=83116968; PubMed=6296778;
RA   Gilson E., Nikaido H., Hofnung M.;
RT   "Sequence of the malK gene in E.coli K12.";
RL   Nucleic Acids Res. 10:7449-7458(1982).
RN   [2]
RP   REVISIONS.
RX   MEDLINE=89384443; PubMed=2674653;
RA   Dahl M.K., Francoz E., Saurin W., Boos W., Manson M.D., Hofnung M.;
RT   "Comparison of sequences from the malB regions of Salmonella
RT   typhimurium and Enterobacter aerogenes with Escherichia coli K12: a
RT   potential new regulatory site in the interoperonic region.";
RL   Mol. Gen. Genet. 218:199-207(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94089392; PubMed=8265357;
RA   Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA   Daniels D.L.;
RT   "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT   region from 89.2 to 92.8 minutes.";
RL   Nucleic Acids Res. 21:5408-5417(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [6]
RP   SEQUENCE OF 1-30 FROM N.A.
RX   MEDLINE=82219202; PubMed=6283312;
RA   Bedouelle H., Hofnung M.;
RT   "A DNA sequence containing the control regions of the malEFG and
RT   malK-lamB operons in Escherichia coli K12.";
RL   Mol. Gen. Genet. 185:82-87(1982).
CC   -!- FUNCTION: THIS IS ONE OF THE FIVE PROTEINS ESSENTIAL TO THE ACTIVE
CC       BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEM FOR MALTOSE AND
CC       MALTODEXTRIN. MALK IS THE TARGET FOR INDUCER EXCLUSION, MEDIATED
CC       BY THE UNPHOSPHORYLATED ENZYME III OF THE PHOSPHOTRANSFERASE
CC       SYSTEM FOR GLUCOSE AND RESULTING IN THE INHIBITION OF MALTOSE
CC       TRANSPORT. MALK HAS ALSO A REGULATORY FUNCTION ON MAL GENE
CC       EXPRESSION.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J01648; AAB59057.1; ALT_SEQ.
DR   EMBL; U00006; AAC43129.1; -.
```

```
DR   EMBL; AE000476; AAC77005.1; -.
DR   EMBL; AE005636; AAG59234.1; -.
DR   EMBL; AP002568; BAB38441.1; -.
DR   EMBL; V00303; CAA23582.1; ALT_INIT.
DR   PIR; A03411; MMECMK.
DR   EcoGene; EG10558; malK.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; Sugar transport; Inner membrane; ATP-binding;
KW   Complete proteome.
FT   NP_BIND      36     43       ATP (BY SIMILARITY).
FT   CONFLICT    102    102       L -> P (IN REF. 1).
SQ   SEQUENCE   371 AA;  40990 MW;  240D58F59450F30E CRC64;


  Query Match             34.1%;  Score 628;  DB 1;  Length 371;
  Best Local Similarity   36.9%;  Pred. No. 4.3e-38;
  Matches  132;  Conservative   77;  Mismatches  117;  Indels   32;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: ::: |  | : :| | :::|: | :|     :| |||||:|||||||  :|  |:
Db       1 MASVQLQNVTKAWGEVVVSKDINLDIHEGEFVVFVGPSGCGKSTLLRMIAGLETITSGDL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           :  : ::|:  |:::| :|||||:||::|||:| :|::|||      || : |: |:  |:
Db      61 FIGEKRMNDTPPAERGVGMVFQSYALYPHLSVAENMSFGLKLAGAKKEVINQRVNQVAEV 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|:|   |||| |||||:||| :   |   ||| |:|||||||:| ||: ||  |  :
Db     121 LQLAHLLDRKPKALSGGQRQRVAIGRTLVAEPSVFLLDEPLSNLDAALRVQMRIEISRLH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
             :| | :|||||| ||| ::|:| |:  | : |:|:| |||| ||: ||| |||  : |
Db     181 KRLGRTMIYVTHDQVEAMTLADKIVVLDAGRVAQVGKPLELYHYPADRFVAGFIGSPKMN 240

Qy     239 IIPANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDS 298
           :|                  :   |:|||: : : :  | : ::  || |:: :  :
Db     241 FLPVK----------------VTATAIDQVQVE-LPMPNRQQVWLPVESRDVQ--VGAN 280

Qy     299 VYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLI 356
           : ||:  :: : :  |  |   |    |   | ||  :: |    :        |||:
Db     281 MSLGIRPEHLLPSDIADVILEGEVQVVE---QLGNETQIHIQIPSIR-------QNLV 328


RESULT   9
AFUC_HAEIN
ID   AFUC_HAEIN     STANDARD;      PRT;   328 AA.
AC   P44531;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ferric transport ATP-binding protein afuC.
GN   AFUC OR HI0126.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR IRON. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32698; AAC21800.1; -.
DR   TIGR; HI0126; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Iron transport; Transport; ATP-binding; Complete proteome.
FT   NP_BIND      39     46       ATP (POTENTIAL).
SQ   SEQUENCE   328 AA;  36046 MW;  59B5298825D3597E CRC64;


  Query Match             34.0%;  Score 627.5;  DB 1;  Length 328;
  Best Local Similarity   41.1%;  Pred. No. 4e-38;
  Matches  134;  Conservative   58;  Mismatches  107;  Indels   27;  Gaps    4;

Qy       8 NAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKI 67
           |  | :    ||:||::|| :|:: |||| |||||||:||::||  :   |: : |   :
Db      11 NITKAFGKAVVIDNLDLTIKRGTMVTLLGPSGCGKTTVLRLVAGLENPTSGQIFIDGEDV 70

Qy      68 NNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYA 127
              |:| :|||:||:|||::: ||| :||  : : |||  |:  : |||: :| :
Db      71 TKSSIQNRDICIVFQSYALFPHMSIGDNVGYGLKMQGIGKEERAQRVKEALELVDLAGFE 130

Qy     128 DRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITT 187
           ||  |::|||||||| || || : | ||| |||||||:| ||  ||: |||:|  :|||:
Db     131 DRFVDQISGGQQQRVALARALVLKPKVLLFDEPLSNLDANLRRSMREKIRELQQRLGITS 190

Qy     188 VYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEK- 246
           :|||||| || |:|||: ||  | | |    |||| :| : |:| |:| ::|    ||
Db     191 LYVTHDQTEAFAVSDEVIVMNKGKIMQKAPAKELYLRPNSLFLANFMGESSIFDGKLENG 250

Qy     247 ----------RSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEE-FIKDESGDIEG-T 294
                     | |     ||  |              | ||||  ::  |  | :
Db     251 VADINGYSVPLKDAAQFNLPDGECL-------------VGIRPEAIYLAAEGSDAQRCE 296

Qy     295 IRDSVYLGLNTDYFIETGFASKIQVS 320
           |: :||:|     :::||    : |:
Db     297 IKSAVYMGQPLGKWLQTGRGKDLLVN 322


RESULT  10
POTG_ECOLI
ID   POTG_ECOLI     STANDARD;      PRT;   377 AA.
AC   P31134;
DT   01-JUL-1993 (Rel. 26, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putrescine transport ATP-binding protein potG.
GN   POTG OR B0855.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93106992; PubMed=8416922;
RA   Pistocchi R., Kashiwagi K., Miyamoto S., Nukui E., Sadakata Y.,
RA   Kobayashi H., Igarashi K.;
RT   "Characteristics of the operon for a putrescine transport system that
RT   maps at 19 minutes on the Escherichia coli chromosome.";
RL   J. Biol. Chem. 268:146-152(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
CC   -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       FOR PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M93239; AAA24410.1; ALT_INIT.
DR   EMBL; AE000187; AAC73942.1; ALT_INIT.
DR   EMBL; D90723; BAA35566.1; ALT_INIT.
DR   PIR; B45313; B45313.
DR   HSSP; P13569; 1NBD.
DR   EcoGene; EG11630; potG.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; ATP-binding; Inner membrane; Complete proteome.
FT   NP_BIND      52     59       ATP (BY SIMILARITY).
SQ   SEQUENCE   377 AA;  41930 MW;  4C13E389C03CF76C CRC64;


  Query Match             34.0%;  Score 627;  DB 1;  Length 377;
  Best Local Similarity   39.0%;  Pred. No. 5.2e-38;
  Matches  135;  Conservative   70;  Mismatches  129;  Indels   12;  Gaps    5;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           ::| |  | |     ::::::|| || :| |||||||||:|||||:|||     |:   |
Db      20 LEIRNLTKSYDGQHAVDDVSLTIYKGEIFALLGASGCGKSTLLRMLAGFEQPSAGQIMLD 79

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
                :: : |  | | |:||:||:|||:||  |:|||| | |:|| |:  : |: | |: :
Db      80 GVDLSQVPPYLRPINMMFQSYALFPHMTVEQNIAFGLKQDKLPKAEIASRVNEMLGLVHM 139

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
           ::| ||| :|||||:||| || :|| | :||:|||:  |: |||  |:  : :|   |
Db     140 QEFAKRKPHQLSGGQRQRVALARSLAKRPKLLLLDEPMGALDKKLRDRMQLEVVDILERV 199

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           |:| | |||||||||| :: :||:|  |   ||| |:|:|  |   : | |||  |:
Db     200 GVTCVMVTHDQEEAMTMAGRIAIMNRGKFVQIGEPEEIYEHPTTRYSAEFIGSVNVFEGV 259

Qy     244 LEKRSDGAYIVFSDG--YALRMPA-LDQVEEQAIHVSIRPEEFIKDESGDIE------GT 294
           |::| :   :: | |  : |:: |    |:   :||::|||: :  |           |
Db     260 LKERQEDGLVLDSPGLVHPLKVDADASVVDNVPVHVALRPEKIMLCEEPPANGCNFAVGE 319

Qy     295 IRDSVYLGLNTDYFI--ETGFASKIQVSEESTFEEDLQK-GNRIRL 337
           :    |||  : | :  ::|     |:       : |   |: :||
Db     320 VIHIAYLGDLSVYHVRLKSGQMISAQLQNAHRHRKGLPTWGDEVRL 365


RESULT  11
MSMX_BACSU
ID   MSMX_BACSU     STANDARD;      PRT;   365 AA.
AC   P94360;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Probable multiple sugar-binding transport ATP-binding protein msmX.
GN   MSMX.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=97124196; PubMed=8969509;
RA   Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA   Miwa Y., Fujita Y.;
RT   "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT   containing the lic and cel loci, and creation of a 177 kb contig
RT   covering the gnt-sacXY region.";
RL   Microbiology 142:3113-3123(1996).
CC   -!- FUNCTION: INVOLVED IN A BINDING PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC       RESPONSIBLE FOR THE UPTAKE OF MULTIPLE SUGARS (POTENTIAL).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D83026; BAA11723.1; -.
DR   EMBL; Z99123; CAB15907.1; -.
DR   SubtiList; BG11954; msmX.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Transport; Sugar transport; Membrane; ATP-binding; Complete proteome.
FT   NP_BIND      37     44       ATP (POTENTIAL).
SQ   SEQUENCE   365 AA;  41366 MW;  C50689EE0E060D32 CRC64;


  Query Match             33.2%;  Score 611.5;  DB 1;  Length 365;
  Best Local Similarity   37.4%;  Pred. No. 6.5e-37;
  Matches  138;  Conservative   66;  Mismatches  130;  Indels   35;  Gaps    8;

Qy       1 MSEIKIINAKKIY-HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
           |:|::: :  | |    | ::: |: |        :| |||||:| |||:||   |  |:
```

```
Db       1 MAELRMEHIYKFYDQKEPAVDDFNLHIADKEFIVFVGPSGCGKSTTLRMVAGLEEISKGD 60

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
           ||  :  ::|:: |  |:| |||||||::||:|| ||:|||| :|:|| |: ::  :  :
Db      61 FYIEGKRVNDVAPKDRDIAMVFQNYALYPHMTVYDNIAFGLKLRKMPKPEIKKRVEEAAK 120

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :: : :|  |||  |||||:||| |  |:  :  | ||||||||||:|||: ||  | ::
Db     121 ILGLEEYLHRKPKALSGGQRQRVALGRAIVRDAKVFLMDEPLSNLDAKLRVQMRAEIIKL 180

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RT 237
             :  ||:|||||| ||: :: :| ||||| ||||| ||::|  | | ||  |||
Db     181 HQRLQTTTIYVTHDQTEALTMATRIVVMKDGKIQQIGTPKDVYEFPENVFVGGFIGSPAM 240

Qy     238 NIIPANLEKRSDGAYIVFSDGYALRMPALDQVEEQA-----IHVSIRPEEFIKDE----- 287
           |    |   :||   : |    :   :  : |:      :   ||||: | ||
Db     241 NFFKGKL---TDGLIKIGSAALTVPEGKMKVLREKGYIGKEVIFGIRPED-IHDELIVVE 296

Qy     288 ---SGDIEGTIRDSVYLGL---------NTDYFIETGFASKIQVSEESTFEEDLQKGN-- 333
             :  |:  |  :  ||          | |:         ||  :| |   |: ||:
Db     297 SYKNSSIKAKINVAELLGSEIMIYSQIDNQDFIARIDARLDIQSGDELTVAFDMNKGHFF 356

Qy     334 ----RIRLR 338
               :|:|
Db     357 DSETEVRIR 365
```

```
RESULT  12
YCJV_ECOLI
ID   YCJV_ECOLI     STANDARD;      PRT;   322 AA.
AC   P77481; P76842;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical ABC transporter ATP-binding protein ycjV.
GN   YCJV OR B1318.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=98389658; PubMed=9721282;
RA   Fajardo D.A., Cheung J., Ito C., Sugawara E., Nikaido H.,
RA   Misra R.;
RT   "Biochemistry and regulation of a novel Escherichia coli K-12 porin
RT   protein, OmpG, which produces unusually large channels.";
RL   J. Bacteriol. 180:4452-4459(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
```

```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U49400; AAC34719.1; -.
DR   EMBL; AE000229; AAC74400.1; -.
DR   EMBL; D90769; BAA14893.1; -.
DR   EMBL; D90770; BAA14900.1; -.
DR   EMBL; D90771; BAA14911.1; -.
DR   EcoGene; EG13919; ycjV.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT   NP_BIND      37     44       ATP (POTENTIAL).
FT   CONFLICT     37     38       GP -> AA (IN REF. 1).
SQ   SEQUENCE   322 AA;  35766 MW;  D130E621FDCEC978 CRC64;
```

```
  Query Match             32.7%;  Score 602.5;  DB 1;  Length 322;
  Best Local Similarity   39.3%;  Pred. No. 2.4e-36;
  Matches  130;  Conservative   67;  Mismatches   99;  Indels   35;  Gaps    7;

Qy       1 MSEIKIINAKKIY-HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
             |::: : : :||| : | |::: |: |        :| |||||:| ||||||   | ||:
Db       1 MAQLSLQHIQKIYDNQVHVVKDFNLEIADKEFIVFVGPSGCGKSTTLRMIAGLEEISGGD 60

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
              |  ::|::    ||| |||||||::||:|| ||:|||| :|: || : :: |   :
Db      61 LLIDGKRMNDVPAKARNIAMVFQNYALYPHMTVYDNMAFGLKMQKIAKEVIDERVNWAAQ 120

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :: : :|  |||  |||||:||| |  |:     | |||||||||:|||: ||  | ::
Db     121 ILGLREYLKRKPGALSGGQRQRVALGRAIVREAGVFLMDEPLSNLDAKLRVQMRAEISKL 180

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RT 237
             ::  | :|||||| ||| :: :| :||||::||:| || :|::||| ||: |||
Db     181 HQKLNTTMIYVTHDQTEAMTMATRIVIMKDGIVQQVGAPKTVYNQPANMFVSGFIGSPAM 240

Qy     238 NIIPANLEKRSDGAYIVFSDGYALRMP----ALDQVEE---QAIHVSIRPEEFIKDESGD 290
           | |   :    ||   | ::   | :|    |: : :|   : | : ||||:
Db     241 NFIRGTI----DGDKFV-TETLKLTIPEEKLAVLKTQESLHKPIVMGIRPED-------- 287

Qy     291 IEGTIRDSVYLGLNTDYFIETGFASKIQVSE 321
                     :: |   |   ::|| |:|
Db     288 -----------IHPDAQEENNISAKISVAE 306
```

```
RESULT  13
CYSA_SYNY3
ID   CYSA_SYNY3     STANDARD;      PRT;   355 AA.
AC   P74548;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sulfate transport ATP-binding protein cysA.
GN   CYSA OR SLR1455.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
```

```
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- FUNCTION: INVOLVED IN THE TRANSPORT OF MULTIPLE SULFUR-CONTAINING
CC       COMPOUNDS, INCLUDING SULFATE AND THIOSULFATE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90916; BAA18655.1; -.
DR   HSSP; P13569; 1NBD.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Sulfate transport; Transport; ATP-binding; Inner membrane;
KW   Complete proteome.
FT   NP_BIND      35     42       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  39195 MW;  34CB746CD68E7EB7 CRC64;
```

```
  Query Match             32.6%;  Score 600;  DB 1;  Length 355;
  Best Local Similarity   37.5%;  Pred. No. 4.2e-36;
  Matches  133;  Conservative   65;  Mismatches  117;  Indels   40;  Gaps    4;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           | | |  | : |   ::::|: :| | |  ||| || ||:|||| |||    : |:   :
Db       3 IIINNVSKQFGDFTALKDINLEVPDGKLVALLGPSGSGKSTLLRAIAGLEEPDQGQIIIN 62

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
             :::  ||||| |||:||:| |||:| |:|||| :| |  :  ::  : | |:|:
Db      63 GQDATHVDIRKRNIGFVFQHYALFKHLTIRQNIAFGLEIRKHPPAKTKERVEELLSLIQL 122

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
             :| | :|||||:||| || |||| | |||:|||   |:||:| ::|  :|::  ||
Db     123 EGLGNRYPSQLSGGQRQRVALARALAVQPQVLLLDEPFGALDAKVRKELRAWLRKLHDEV 182

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           :|:|:|||||||||| ::|:| || :| |:|:|  :|:|  ||: ||  |||  |::| |
Db     183 HLTSVFVTHDQEEAMEVADEIVVMSNGKIEQVGTAEEIYEHPASPFVMGFIGEVNVLPRN 242

Qy     244 LEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEF---IKDESGDIEGTIRDSVY 300
             :  |:   |     : |           | :|| :|    : :   : |||:  ::
Db     243 ASLFNYHAFEPHSSNNGHQEP----------VFVRPHDFELLTEADDASVAGTIKRVIH 291

Qy     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNL 355
           ||            |:|||               : |  ||  |   : :  | |
Db     292 LG----------SEIQV-------------EVLLMDNTAVLAYLNREQGQQL 320
```

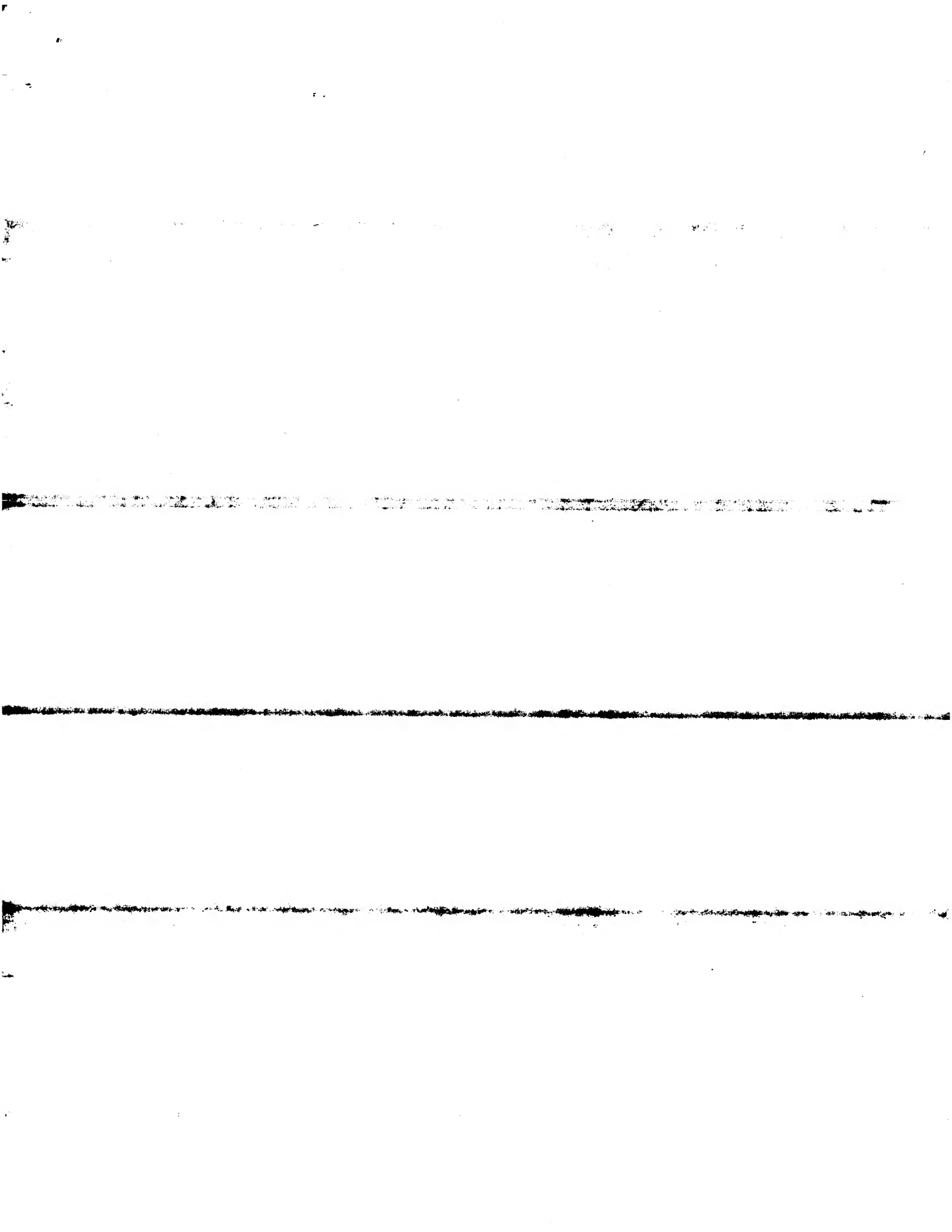
```
RESULT  14
SMOK_RHOSH
ID   SMOK_RHOSH     STANDARD;      PRT;   332 AA.
AC   P54933; O30834;
DT   01-OCT-1996 (Rel. 34, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   ATP-binding transport protein smoK (PolK).
GN   SMOK OR POLK.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
```

```
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SI-4;
RX   MEDLINE=97474256; PubMed=9335280;
RA   Stein M.A., Schafer A., Giffhorn F.;
RT   "Cloning, nucleotide sequence, and overexpression of smoS, a
RT   component of a novel operon encoding an ABC transporter and polyol
RT   dehydrogenases of Rhodobacter sphaeroides Si4.";
RL   J. Bacteriol. 179:6335-6340(1997).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF018073; AAC45769.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport.
FT   NP_BIND      36     43       ATP (POTENTIAL).
SQ   SEQUENCE   332 AA;  36362 MW;  36B26464A707C1B4 CRC64;
```

```
  Query Match             32.5%;  Score 599;  DB 1;  Length 332;
  Best Local Similarity   36.3%;  Pred. No. 4.5e-36;
  Matches  130;  Conservative   67;  Mismatches  125;  Indels   36;  Gaps    5;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           | :| : | :| : :  || :|:: |  |     :| ||||:||||:|||   :  |:
Db       1 MGKITLRNVQKRFGEAVVIPSLDLDIEDGEFVVFVGPSGCGKSTLLRLIAGLEDVSDGQI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             |      | |:|| : ||||:||::||:||: |:|| |   |:  :|: :: :   ::
Db      61 MIDGRDATEMPPAKRGLAMVFQSYALYPHMTVKKNIAFPLRMAKMEPQEIERRVSNAAKI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
             : :  | ||:| :|||||:||| :  |:   |:  | |||||||:| ||::||  | |:
Db     121 LNLTNYLDRRPGQLSGGQRQRVAIGRAIVREPAAFLFDEPLSNLDAALRVNMRLEITELH 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :  | :|||||| ||| ::|:| |:  | |:|:| |  ||  ||| ||| |||
Db     181 QSLETTMIYVTHDQVEAMTMADKIVVLNAGRIEQVGSPLTLYRNPANLFVAGFIG----- 235

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVE--EQAIH----VSIRPEEF-IKDESGDIEG 293
                                 | :: :|  | | |    : ||||   :  |:|  ||
Db     236 ----------------------SPKMNLIEGPEAAKHGATTIGIRPEHIDLSREAGAWEG 273

Qy     294 TIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
             :  | :||  :| |:    |    ::  :  |  :  |:|: |     |:: | |||
Db     274 EVGVSEHLG--SDTFLHVHVAGMPTLTVRTGGEFGVHHGDRVWLTPQADKIHRFGADG 329
```

```
RESULT  15
Y4OS_RHISN
ID   Y4OS_RHISN     STANDARD;      PRT;   371 AA.
AC   P55604;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable ABC transporter ATP-binding protein Y4OS.
GN   Y4OS.
OS   Rhizobium sp. (strain NGR234).
```

```
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
CC   -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
CC       SYSTEM Y4OPQRS. THIS SYSTEM PROBABLY TRANSPORTS A SUGAR-LIKE
CC       MOLECULE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC       TRANSPORT SYSTEM.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000089; AAB91805.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   Hypothetical protein; ATP-binding; Transport; Inner membrane; Plasmid.
FT   NP_BIND      52     59       ATP (POTENTIAL).
SQ   SEQUENCE   371 AA;  40580 MW;  0F0CFFC186D516C2 CRC64;


  Query Match             32.1%;  Score 592;  DB 1;  Length 371;
  Best Local Similarity   38.8%;  Pred. No. 1.7e-35;
  Matches  135;  Conservative   63;  Mismatches  126;  Indels   24;  Gaps    7;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |::: | |  | |  : ||  |:  |  |    |:| |||||:|||||:||   | ||:
Db      17 MTDVTIRNVTKRYGALTVIPQLSFRIEDGEFVVLVGPSGCGKSTLLRMLAGLEEISGGDL 76

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
                ||:    :|:: :||||||::||:|| :|: | |  :| |: |: :: :|   :
Db      77 LMGADVINDRPAKERDMAIVFQNYALYPHMTVAENMGFALKLRKRPRAEIDERVDKAAAI 136

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           : : :  || |  |||||:||| :  |:  :| | | |||||||:||||: ||  |: :
Db     137 LGLGKLLDRYPRALSGGQRQRVAMGRAIVRDPQVFLFDEPLSNLDAKLRVQMRAEIKALH 196

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
             ||||||||| ||| ::|:| || :| ::|:| | ||| :||| ||| |||    |
Db     197 QRLKITTVYVTHDQIEAMTMADKIVVMNEGRVEQMGTPLELYDRPANIFVAGFIGSPSMN 256

Qy     239 IIPANLEKRSDGAYIVFSDGYALRM---PALDQVEEQAIHVSIRPEEFIKDESG------ 289
            :|| :   ::|  :   :| || :   | |    |  :   ||||    |:|
Db     257 FLPATV-AATNGPLLKTPEGVALPIDGGPTLTGRSE--VTYGIRPEHLQLGETGIPAEVV 313

Qy     290 DIEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRL 337
           :| |\       |  |  ::  |    : |  :     |:: | :| |
Db     314 VVEPT-------GSETQLYVTVGGREVVAVLRDRV---DVRPGEKIWL 351


Search completed: July 30, 2002, 15:51:54
Job time: 222 sec
```



GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 30, 2002, 15:47:47 ; Search time 28.53 Seconds
(without alignments)
2201.092 Million cell updates/sec

Title: US-09-769-787-162
Perfect score: 1843
Sequence: 1 MSEIKIINAKKIYHDVPVIE.....LNIFSADGSQNLIKGVNHGT 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1843	100.0	363	16	Q97ST2	Q97st2 streptococc
2	919	49.9	374	2	O54370	O54370 treponema h
3	714.5	38.8	355	16	Q9X0V9	Q9x0v9 thermotoga
4	701	38.0	368	16	Q9X196	Q9x196 thermotoga
5	688.5	37.4	337	1	Q9HH28	Q9hh28 pyrococcus
6	683.5	37.1	337	1	Q9HH23	Q9hh23 thermococcu
7	672	36.5	353	16	Q92NI2	Q92ni2 rhizobium m
8	670.5	36.4	358	17	Q9YB65	Q9yb65 aeropyrum p
9	668	36.2	373	2	Q9L531	Q91531 vibrio chol
10	668	36.2	373	16	Q9KL04	Q9kl04 vibrio chol
11	662.5	35.9	372	17	Q9HKB3	Q9hkb3 thermoplasm
12	661.5	35.9	397	2	O86831	O86831 streptomyce
13	661	35.9	375	17	O57933	O57933 pyrococcus
14	658.5	35.7	392	16	Q49978	Q49978 mycobacteri
15	657.5	35.7	369	16	Q9X103	Q9x103 thermotoga
16	656	35.6	364	16	Q99V03	Q99v03 staphylococ
17	655.5	35.6	352	16	Q9CKZ6	Q9ckz6 pasteurella
18	655.5	35.6	359	16	Q92NG2	Q92ng2 rhizobium m
19	653	35.4	358	16	Q92XF6	Q92xf6 rhizobium m
20	651	35.3	364	17	Q9V2Q1	Q9v2q1 pyrococcus
21	650.5	35.3	347	16	O51587	O51587 borrelia bu
22	650	35.3	373	17	O57758	O57758 pyrococcus
23	649.5	35.2	357	17	O58504	O58504 pyrococcus
24	649	35.2	378	16	Q9KDY5	Q9kdy5 bacillus ha
25	648	35.2	381	17	Q9YG58	Q9yg58 aeropyrum p
26	645.5	35.0	371	17	Q9V293	Q9v293 pyrococcus
27	645	35.0	370	16	Q92L00	Q92l00 rhizobium m
28	644.5	35.0	329	16	Q9PPH5	Q9pph5 campylobact
29	644.5	35.0	364	16	Q98HF7	Q98hf7 rhizobium l
30	643.5	34.9	370	16	Q9I6T2	Q9i6t2 pseudomonas
31	643.5	34.9	381	16	Q9A3R8	Q9a3r8 caulobacter
32	642.5	34.9	369	2	Q9LBW0	Q9lbw0 clostridium
33	642.5	34.9	408	16	P73468	P73468 synechocyst
34	641.5	34.8	377	16	Q9KS33	Q9ks33 vibrio chol
35	641	34.8	370	16	Q98CK7	Q98ck7 rhizobium l
36	639.5	34.7	380	16	Q9CP06	Q9cp06 pasteurella
37	637.5	34.6	362	16	Q92UQ0	Q92uq0 rhizobium m
38	637	34.6	355	16	Q92TX6	Q92tx6 rhizobium m
39	636.5	34.5	359	16	Q9WYQ2	Q9wyq2 thermotoga
40	636.5	34.5	372	1	Q9HH32	Q9hh32 pyrococcus
41	636	34.5	377	2	P96483	P96483 streptomyce
42	635.5	34.5	372	1	Q9V306	Q9v306 thermococcu
43	634.5	34.4	362	16	Q9RUT0	Q9rut0 deinococcus
44	634.5	34.4	372	16	Q92ZG9	Q92zg9 rhizobium m
45	634.5	34.4	380	16	Q988Z0	Q988z0 rhizobium l

ALIGNMENTS

```
RESULT 1
Q97ST2
ID   Q97ST2      PRELIMINARY;      PRT;   363 AA.
AC   Q97ST2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   SP0242.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007337; AAK74421.1; -.
DR   TIGR; SP0242; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   363 AA;  40757 MW;  3D9F10B1D8DD7731 CRC64;
```

```
  Query Match             100.0%;  Score 1843;  DB 16;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-115;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVY 300

Qy     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LGLNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNLIKGVN 360

Qy     361 HGT 363
           |||
Db     361 HGT 363


RESULT   2
O54370
ID   O54370      PRELIMINARY;      PRT;   374 AA.
AC   O54370;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ABC TRANSPORTER BITD.
GN   BIT OPERON.
OS   Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC   Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
OX   NCBI_TaxID=159;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20026807; PubMed=10559160;
RA   Dugourd D., Martin C., Rioux C.R., Jacques M., Harel J.;
RT   "Characterization of a periplasmic ATP-binding cassette iron import
RT   system of Brachyspira (Serpulina) hyodysenteriae.";
RL   J. Bacteriol. 181:6948-6957(1999).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR   EMBL; U75349; AAB95372.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport.
SQ   SEQUENCE   374 AA;  42638 MW;  E5364E50A2A39E93 CRC64;


  Query Match             49.9%;  Score 919;  DB 2;  Length 374;
  Best Local Similarity   51.4%;  Pred. No. 1.3e-53;
  Matches  188;  Conservative   65;  Mismatches  103;  Indels   10;  Gaps    4;

Qy       4 IKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFD 63
           | | |  | |  : :| :|:: |  |  ||||| ||||||||||||||||:|||||  ||
Db       5 ISIENVVKRYEKLTIIPDLSLEIKNGEFFTLLGPSGCGKTTLLRMIAGFNTIEGGEIKFD 64

Qy      64 DTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQI 123
              |||:  ||||||||||||||||:|||:|| :||  :|  || : :: :: | :::|
Db      65 KDVINNIPAHKRNIGMVFQNYAIFPHMTVRENVEYGLKLRKENKESMKKKVDEMLHVVKI 124

Qy     124 AQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEV 183
            :| || |::|||||||| || |: : |||||||||||||:||||::|| ||::||  |
Db     125 EEYQDRLPERLSGGQQQRVALARAIVITPSVLLMDEPLSNLDAKLRIEMRSAIKDIQRHV 184

Qy     184 GITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPAN 243
           |||||||||||||:|:||:|||||:|||||:| |  :| :| | ||||||| :|:  |
Db     185 GITTVYVTHDQEEALAVSDRIAVMKNGVIQQVGSPVSIYTRPYNVFVATFIGHSNLFYAT 244

Qy     244 LEKRSDGAYIVFSDGYALRMP-ALDQVEEQAIHVSIRPEEFIKDESGD--IEGTIRDSVY 300
           ::   :  |::|  || |:|   ||  :   : | ||||||  :   |  |:  |   :
Db     245 IKIEGNDTYLLFRCGYKLKMDNLLDVKDGDEVVVGIRPEEFFVNSENDEGIKAKILSKTF 304

Qy     301 LGLNTDYFIE------TGFASKIQVSEESTFEEDL-QKGNRIRLRINTQKLNIFSADGSQ 353
           ||  |:||:             |: |::|:: : : :|   | |: |  |:|:|: |  :
Db     305 LGKYTNYFLHFNDNEVVPDQPSIEYSQDSSYTDRMYEKDEVITLKPNANKINVFTPDMEK 364

Qy     354 NLIKGV 359
           :|||||
Db     365 SLIKGV 370


RESULT   3
Q9X0V9
ID   Q9X0V9      PRELIMINARY;      PRT;   355 AA.
AC   Q9X0V9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   SUGAR ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   TM1232.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS).
DR   EMBL; AE001779; AAD36307.1; -.
DR   TIGR; TM1232; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Complete proteome; Transport.
SQ   SEQUENCE   355 AA;  39814 MW;  44F330C0316CE3EF CRC64;


  Query Match             38.8%;  Score 714.5;  DB 16;  Length 355;
  Best Local Similarity   39.4%;  Pred. No. 5.2e-40;
  Matches  142;  Conservative   82;  Mismatches  125;  Indels   11;  Gaps    5;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |:::||   || : :|  :: ::: : :|    ||| |||||||||| |||   : ||:
Db       1 MAQVKIDGVKKYFGNVRALDGIDLVVNEGEFLVLLGPSGCGKTTLLRCIAGLEQVTGGKI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
```

```
          :|:|  : |: |  ||| ||||:||::||: | ||:|: |  |||||||: ::     :|
Db     61 FFNDRDVTNLPPKDRNISMVFQSYAVWPHMKVYDNIAYPLKLKKVPKEEIEKRVKWAADL 120

Qy    121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
          : |::  || | :|||||:||| :| |:   | ||||||||||:| ||: ||  ::::|
Db    121 LHISELLDRYPAQLSGGQRQRVAVARAIVHEPEVLLMDEPLSNLDALLRVKMRSELKKLQ 180

Qy    181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
            :|:||:|||||| ||| : |:|||  | :||:| | |:|| | | ||| |:|   :
Db    181 ERIGVTTIYVTHDQTEAMTMGDRIAVMNQGKLQQVGTPSEIYHHPVNIFVAGFVGSPQMN 240

Qy    241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGD---IEGTIRD 297
            :| ||:|  :|  :|  :::||      : : : ||||   :|  :   :||
Db    241 FLEMEVRSEGNSVVLQNG-EIKIPAKTDPGAKKVILGIRPENVYLEEKPNTLKLEG---- 295

Qy    298 SVYLG--LNTDYFIETGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADGSQNL 355
           ||    | :|  :     |: ::  :   : | : | :|   :: :|:::|  :  : :
Db    296 EVYFAEKLMSDTILHLNVGSE-KIVAKIPGDVDFRSGEKITFFLDVEKIHLFHPETGERI 354
```

RESULT 4
Q9X196
ID Q9X196 PRELIMINARY; PRT; 368 AA.
AC Q9X196;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE SPERMIDINE/PUTRESCINE ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN TM1376.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
DR EMBL; AE001791; AAD36446.1; -.
DR TIGR; TM1376; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Complete proteome; Transport.
SQ SEQUENCE 368 AA; 42046 MW; 89954C183F611C9C CRC64;

Query Match 38.0%; Score 701; DB 16; Length 368;
Best Local Similarity 42.2%; Pred. No. 4.3e-39;
Matches 138; Conservative 75; Mismatches 102; Indels 12; Gaps 4;

```
Qy      3 EIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYF 62
          |: | |  | : |  |::|::: | ||  |::|| ||||||||||:|||  :| |:
Db      5 EVSIKNVSKFFDDFQVLKNVSLDIKKGEFFSILGPSGCGKTTLLRVIAGFEGVESGDVLL 64

Qy     63 DDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQ 122
          |   | |: |:|| : ::|||||:|||||| :|:|| |  ||: : |: |: |: | |::
Db     65 DGKSILNLPPNKRPVNIIFQNYALFPHLTVFENIAFPLKLKKLSENEINQRVNELLSLIR 124

Qy    123 IAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHE 182
          : ::| : | :|||||:||| :| |||  | |||:||||| |:|||| ::   :  :
Db    125 MEEHAQKMPSQLSGGQKQRVAIARALANEPRVLLLDEPLSALDAKLRQELLVELDNLHDR 184

Qy    183 VGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPA 242
          |||| :|||||| ||:::||::|:| :| | |:| | |:|  | | | ||||| ||:: |
Db    185 VGITFIYVTHDQAEAISVSDRVALMNEGEIVQVGTPYEVYESPVNVFAATFIGETNLMKA 244

Qy    243 NLEKRSDGAYIVFSDGYALRMPALDQVEEQA--IHVSIRPE-------EFIKDESGDI-E 292
          : :  |  |:| | |         |:  ::   : :::|||       :|  :|: ::
Db    245 EVVEVEDEYYVVESPGIGQFRCYRDKEAKKGDRLLITLRPEKIRISRKQFRSEETFNVFH 304

Qy    293 GTIRDSVYLGLNTDYFI--ETGFASKI 317
          | : : :|:|  | ||:  : |:  |:
Db    305 GVVDEEIYMGHQTKYFVRLDEGYIMKV 331
```

RESULT 5
Q9HH28
ID Q9HH28 PRELIMINARY; PRT; 337 AA.
AC Q9HH28;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE SUGAR-BINDING TRANSPORT ATP-BINDING PROTEIN.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20566786; PubMed=11115105;
RA Diruggiero J., Dunn D., Maeder D.L., Holley-Shanks R., Chatard J.,
RA Horlacher R., Robb F.T., Boos W., Weiss R.B.;
RT "Evidence of recent lateral gene transfer among hyperthermophilic
RT archaea.";
RL Mol. Microbiol. 38:684-693(2000).
DR EMBL; AF307052; AAG45383.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding.
SQ SEQUENCE 337 AA; 38376 MW; E046DFD21B8CF1A0 CRC64;

Query Match 37.4%; Score 688.5; DB 1; Length 337;
Best Local Similarity 44.6%; Pred. No. 2.6e-38;
Matches 149; Conservative 55; Mismatches 71; Indels 59; Gaps 7;

```
Qy     15 DVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEPSK 74
          |:| :|        |   |||| ||||||| ||:||||   : |: |||| :|:: | :
Db     19 DIPFLE-----AKPGEFLTLLGPSGCGKTTTLRIIAGFEKPDKGKIYFDDTVMNDVPPYE 73

Qy     75 RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRKPDKL 134
          ||||:|||:||:|||:|| ||::|||  :|: |||: :: :  |||: :  : || |::|
Db     74 RNIGIVFQDYALFPHMTVYDNISFGLKLRKLSKEEIKRRVSWALELVGLKGFEDRYPEQL 133

Qy    135 SGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITTVYVTHDQ 194
          |||||||| || || : | :||:|||||||:||:|  :|  |: || |:||||:||||||
Db    134 SGGQQQRVALARALVIEPQLLLLDEPLSNLDAKIRERLRGEIKRIQKELGITTIYVTHDQ 193

Qy    195 EEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEKRSDGAYI- 253
          ||||||||:||||  | |:|:| | :||: | ||||| |:| :||:  | |:    ||:
Db    194 EEAMAISDRIAVMSVGKIEQVGNPLDLYYNPKNEFVARFLGLSNILEVNAEEGK--AYVG 251

Qy    254 -----VFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVYLGLNTDYF 308
               | ::|                : :  |||  :  | ||           :|   ||
Db    252 NLCFEVKTEG--------------RVKIFFRPES-VYIEEGD----------MGEIVDYE 286

Qy    309 IETGFASKIQVSEESTFEEDLQKGNRIRLRINTQ 342
          :  |                      ||||:|: :
Db    287 LLPG--------------------RIRLKIDVE 299
```

```
RESULT   6
Q9HH23
ID   Q9HH23      PRELIMINARY;      PRT;   337 AA.
AC   Q9HH23;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE SUGAR-BINDING TRANSPORT ATP-BINDING PROTEIN.
OS   Thermococcus litoralis.
OC   Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX   NCBI_TaxID=2265;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20566786; PubMed=11115105;
RA   Diruggiero J., Dunn D., Maeder D.L., Holley-Shanks R., Chatard J.,
RA   Horlacher R., Robb F.T., Boos W., Weiss R.B.;
RT   "Evidence of recent lateral gene transfer among hyperthermophilic
RT   archaea.";
RL   Mol. Microbiol. 38:684-693(2000).
DR   EMBL; AF307053; AAG45398.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding.
SQ   SEQUENCE   337 AA;  38343 MW;  12DFA3B7252984AA CRC64;


  Query Match             37.1%;  Score 683.5;  DB 1;  Length 337;
  Best Local Similarity   44.6%;  Pred. No. 5.7e-38;
  Matches  149;  Conservative   55;  Mismatches   71;  Indels   59;  Gaps    7;

Qy      15 DVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYFDDTKINNMEPSK 74
           |:| :|       |   |||| ||||||| ||:|||   : |: ||:|| :|:: | :
Db      19 DIPSLE-----AKPGEFLTLLGPSGCGKTTTLRIIAGFEKPDKGKIYFNDTVMNDVPPYE 73

Qy      75 RNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQIAQYADRKPDKL 134
           ||||:|||:||:|||:|| ||::||||  :|: |||: :: :  |||: :  : || |::|
Db      74 RNIGIVFQDYALFPHMTVYDNISFGLKLRKLSKEEIKRRVSWALELVGLKGFEDRYPEQL 133

Qy     135 SGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHEVGITTVYVTHDQ 194
           |||||||| || || : | :||:|||||||:||:|  :|  |: || |:||||:||||||
Db     134 SGGQQQRVALARALVIEPQLLLLDEPLSNLDAKIRERLRGEIKRIQKELGITTIYVTHDQ 193

Qy     195 EEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPANLEKRSDGAYI- 253
           ||||||||:|||| | |:||| | :||: | ||||| |:| :||:  | |:    ||:
Db     194 EEAMAISDRIAVMSVGKIEQIGNPLDLYYNPKNEFVARFLGLSNILEVNAEEGK--AYVG 251

Qy     254 -----VFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDIEGTIRDSVYLGLNTDYF 308
                | ::|               : :  |||  :  | ||          :|   ||
Db     252 NLCFEVKTEG--------------RVKIFFRPES-VYIEEGD----------MGEIVDYE 286

Qy     309 IETGFASKIQVSEESTFEEDLQKGNRIRLRINTQ 342
           :  |                    ||||:|: :
Db     287 LLPG--------------------RIRLKIDVE 299


RESULT   7
Q92NI2
ID   Q92NI2      PRELIMINARY;      PRT;   353 AA.
AC   Q92NI2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ATP-BINDING ABC TRANSPORTER PROTEIN.
GN   SMC01608.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21368234; PubMed=11474104;
RA   Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA   Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA   Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA   Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA   Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA   Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA   Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA   Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA   Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA   Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT   "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL   Science 293:668-672(2001).
DR   EMBL; AL591790; CAC46799.1; -.
KW   Complete proteome.
SQ   SEQUENCE   353 AA;  37903 MW;  09B8D5E8A74ABEDF CRC64;


  Query Match             36.5%;  Score 672;  DB 16;  Length 353;
  Best Local Similarity   43.8%;  Pred. No. 3.5e-37;
  Matches  142;  Conservative   65;  Mismatches  109;  Indels    8;  Gaps    3;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
             |: ::: | :| :  | |: | :: | ||   : || |||||||:|||||| :  ||
Db       1 MAFLQLTNIQKSFGPVQVVHNFDMGIDKGEFVSFLGPSGCGKTTILRMIAGFETPSGGSI 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           : |      ::|::|||||||| ||:||:: | |||||||      : |:  :  : | |
Db      61 FIDGKDQGALKPNQRNIGMVFQAYALFPNMNVHDNVAFGLKVAGASRTEIEARVKQMLGL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :::   ||| | :|||||||||| || ||||  | |||:|||| |:||:|: :|: ||:||
Db     121 IKLEHLADRFPYQLSGGQQQRVALARALAVKPQVLLLDEPLSALDAKIRISLREEIRQIQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
           ::|||||:||||||||::|||:| ||  |   ||| | |:|: ||  |||:|:|  |||
Db     181 QQLGITTVFVTHDQEEALSISDRIVVMNSGRADQIGTPFEIYNTPATRFVASFVGTLNII 240

Qy     241 PANLEKRSDGAYIVFSDGYALRMPALDQVEEQAIHVSIRPEEFIKDESGDI-EGTIRDSV 299
           :   : ||  :    :|: |      :| :::||      |:| | ||   |:
Db     241 EGKVADPASGAVTIGGQRISLKEPIAGAKGGDSISLALRP------EAGSIAEGARGDTA 294

Qy     300 YLG-LNTDYFIETGFASKIQVSEE 322
             | : :  |: :   :|::| ::
Db     295 LPGEVVSRSFLGSVIRTKLRVGKD 318


RESULT   8
Q9YB65
ID   Q9YB65      PRELIMINARY;      PRT;   358 AA.
AC   Q9YB65;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   358AA LONG HYPOTHETICAL TRANSPORTER ATP-BINDING PROTEIN.
GN   APE1732.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
```

```
RL   DNA Res. 6:83-101(1999).
CC   -!- COFACTOR: REQUIRES ZINC FOR ITS ACTIVITY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS).
DR   EMBL; AP000062; BAA80733.1; -.
DR   HSSP; Q58663; 1G6H.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR002328; ADH_zinc.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR   PROSITE; PS00059; ADH_ZINC; 1.
KW   ATP-binding; Complete proteome; Oxidoreductase; Transport; Zinc.
SQ   SEQUENCE   358 AA;  40173 MW;  833B0813EA9971D7 CRC64;


  Query Match             36.4%;  Score 670.5;  DB 17;  Length 358;
  Best Local Similarity   39.6%;  Pred. No. 4.5e-37;
  Matches  145;  Conservative   76;  Mismatches  122;  Indels   23;  Gaps    8;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: |:: :  | : :   :: ::: |  | :||||| ||||||| ||:|||| : : |
Db       1 MAGIRLESVTKRFGNTVALDRVSLNISDGEIFTLLGPSGCGKTTTLRVIAGFETPDEGRV 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           |      :  ::| :||  |||||||::||: | ||:|:||  :|:|: |::::     ||
Db      61 YIGSRDVTMLKPYERNTAMVFQNYALWPHMRVFDNIAYGLKLRKLPRSEIVRRVRWAAEL 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           ::|    || | :|||||||||| :| |:   | ||||||||||:| ||| ||: |  :|
Db     121 LEIDHLLDRYPHQLSGGQQQRVAVARAIVTEPEVLLMDEPLSNLDAHLRLKMREEIVRLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNII 240
             :|:| ||||||||||::|| ::|||  | ::|:| | |:| |||  |||||||: ::
Db     181 KRLGVTIVYVTHDQEEALSISHRVAVMNRGRVEQVGTPMEVYEKPATYFVATFIGRSTVL 240

Qy     241 PANL-EKRSDGAYIVFSDGYALRMPALDQV----EEQAIHVSIRPEEFIK---DESGD-- 290
           : |    |   | :|  | :   |      : : : | ||||  :|   : :|:
Db     241 DGRVSEVLGSGMVRVALEG-GLSIVGTDMEGGLRQGERVKVVIRPER-VKVGHEYNGENV 298

Qy     291 IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEEDLQK----GNRIRLRINTQKLNI 346
           || :  :::||  |   :| |        :| |   | ::    |  :|  |: :: :
Db     299 FEGKVSLAMFLGWRTQLKVEVG-------GQEITIYSDPRRAPLPGQPVRFYIDPEEAKV 351

Qy     347 FSADGS 352
           :   ||
Db     352 YRQAGS 357


RESULT   9
Q9L531
ID   Q9L531      PRELIMINARY;      PRT;   373 AA.
AC   Q9L531;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MALK.
GN   MALK.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=569B;
RA   Dutta P.P., Roychoudhury S., Chaudhuri K.;
RT   "Probable MalK protein of Vibrio cholerae.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF259974; AAF70304.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   ATP-binding.
SQ   SEQUENCE   373 AA;  41404 MW;  BDAC521894CABA55 CRC64;


  Query Match             36.2%;  Score 668;  DB 2;  Length 373;
  Best Local Similarity   38.4%;  Pred. No. 7e-37;
  Matches  138;  Conservative   76;  Mismatches  135;  Indels   10;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: : : |  | | || : :|::: | :|     :| |||||:|||| |||   |  |:
Db       1 MASVTLKNVCKAYGDVLISKNVDLEIQEGEFVVFVGPSGCGKSTLLRCIAGLEDITSGDL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           |  : ::|::||||| :|||||:||::||| : ||::|||   |  | |: :: :   |:
Db      61 YIGEQRMNDVEPSKRGVGMVFQSYALYPHLNLYDNMSFGLKLSKADKSEIKKRVDHAAEI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|::   ||:|  |||||:||| :   |   |:| |:|||||||:| ||: ||  | ::|
Db     121 LQLSHLLDRQPKALSGGQRQRVAIGRTLVSQPNVFLLDEPLSNLDAALRVQMRSEITKLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
           ::| | :|||||| ||| ::|:|  |:   | : |:|:| |||| | | ||| |||  : |
Db     181 RKLGCTMIYVTHDQVEAMTMADKIVVLDAGFVSQVGKPLELYHYPQNRFVAGFIGSPKMN 240

Qy     239 IIPANLE-KRSDGAYIVFSDGYALRMP--ALDQVEEQAIHVSIRPEEFIKDESGD--IEG 293
           :   :|    |   :  |:|    :|          : : : |||  :: | ||  |||
Db     241 FMSVFIEGVEKDRVQVQLSNGTTFWIPVDGTTVTRGERMSLGIRPEHLVEAEHGDAKIEG 300

Qy     294 TIRDSVYLGLNTDYFIE-TGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           :     ||  |  ::   |  | :   :  |   |:: |: : : |   : ::| :||
Db     301 KVMIVEKLGNETQVYMNLKGSDSDVIYRQPDTL--DVETGDTLTIGIPAHRCHLFHSDG 357


RESULT  10
Q9KL04
ID   Q9KL04      PRELIMINARY;      PRT;   373 AA.
AC   Q9KL04;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MALTOSE/MALTODEXTRIN ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   VCA0946.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EL TOR N16961 / SEROTYPE O1;
RX   MEDLINE=20406833; PubMed=10952301;
RA   Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA   Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA   Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA   Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA   McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA   Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA   Fraser C.M.;
RT   "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT   cholerae.";
RL   Nature 406:477-483(2000).
DR   EMBL; AE004421; AAF96842.1; -.
DR   TIGR; VCA0946; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   373 AA;  41404 MW;  A15E521894CABA55 CRC64;
```

```
  Query Match           36.2%;  Score 668;  DB 16;  Length 373;
  Best Local Similarity  38.4%;  Pred. No. 7e-37;
  Matches  138;  Conservative   76;  Mismatches  135;  Indels   10;  Gaps    6;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           |: : : |  | | || : :|::: | :|     :| |||||:|||| |||   |  |:
Db       1 MASVTLKNVCKAYGDVLISKNVDLEIQEGEFVVFVGPSGCGKSTLLRCIAGLEDITSGDL 60

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
           |  : ::|::|||| :||||:||::||| : ||::|||   |  | |: :: :   |:
Db      61 YIGEQRMNDVEPSKRGVGMVFQSYALYPHLNLYDNMSFGLKLSKADKSEIKKRVDHAAEI 120

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           :|::   ||:|  |||||:||| :   |   |:| |:|||||||:| ||: ||  | ::|
Db     121 LQLSHLLDRQPKALSGGQRQRVAIGRTLVSQPNVFLLDEPLSNLDAALRVQMRSEITKLQ 180

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIG--RTN 238
           ::| | :|||||| ||| ::|:| |:  | : |:|:| |||| | | ||| |||  : |
Db     181 RKLGCTMIYVTHDQVEAMTMADKIVVLDAGFVSQVGKPLELYHYPQNRFVAGFIGSPKMN 240

Qy     239 IIPANLE-KRSDGAYIVFSDGYALRMP--ALDQVEEQAIHVSIRPEEFIKDESGD--IEG 293
           :   :|    |   :  |:|    :|        : : : |||| :: | ||  |||
Db     241 FMSVFIEGVEKDRVQVQLSNGTTFWIPVDGTTVTRGERMSLGIRPEHLVEAEHGDAKIEG 300

Qy     294 TIRDSVYLGLNTDYFIE-TGFASKIQVSEESTFEEDLQKGNRIRLRINTQKLNIFSADG 351
           :     ||  |  ::   |  | :   :  |   |:: |: : : |   : ::| :||
Db     301 KVMIVEKLGNETQVYMNLKGSDSDVIYRQPDTL--DVETGDTLTIGIPAHRCHLFHSDG 357


RESULT  11
Q9HKB3
ID   Q9HKB3      PRELIMINARY;      PRT;   372 AA.
AC   Q9HKB3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PROBABLE SUGAR ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   TA0688.
OS   Thermoplasma acidophilum.
OC   Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC   Thermoplasma.
OX   NCBI_TaxID=2303;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 1728;
RX   MEDLINE=20479972; PubMed=11029001;
RA   Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA   Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT   "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT   acidophilum.";
RL   Nature 407:508-513(2000).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS).
DR   EMBL; AL445065; CAC11826.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Complete proteome; Transport.
SQ   SEQUENCE   372 AA;  41642 MW;  2F37EC38218AE4B9 CRC64;


  Query Match           35.9%;  Score 662.5;  DB 17;  Length 372;
  Best Local Similarity  39.3%;  Pred. No. 1.6e-36;
  Matches  149;  Conservative   72;  Mismatches  125;  Indels   33;  Gaps    9;

Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
           || | | |    | |  :::| |: :  |  | :||:|| ||||||| |||   |: |
Db       1 MSVI-IENLTARYGDKYILKNFNLRVEDGEFFVILGSSGSGKTTLLRSIAGLIPIDDGRI 59

Qy      61 YFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLEL 120
             |   : :: || ||| |||||:|::||:|| ||::  |  | |||:|: :: :   |:
Db      60 LIDGNDVTDLYPSDRNIAMVFQNFALYPHMTVYDNISLNLKIKHVPKDEIKKKVSDVAEM 119

Qy     121 MQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQ 180
           : |:|:  : | :||||:|||| :| |:  :||| ||||||||:|||| :|   :|
Db     120 LHISQHLQKYPRQLSGGEQQRVGIARAMVRDPSVFLMDEPLSNLDAKLRREMLGELRSFH 179

Qy     181 HEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR--TN 238
           :|  | :|| |||:||||::|:| |:  | | |:| | |||  ||| ||| |||    |
Db     180 EKVKKTIIYVCHDQDEAMALADRILVLNQGQIMQLGTPDELYDHPANVFVAGFIGNPPMN 239

Qy     239 IIPANLEKRSDGAYI-VFSDGYALRMPALDQVEEQAIHVSIRPEEF-IKDESGDIEGTIR 296
           |:  : |::| |  | | : ||   :   |:: |  : : |||   : |  |
Db     240 ILECDAEEKSGGTRIDVMNSGY---LEVDDEIPEGKMFIGIRPELIGLSDHDG------- 289

Qy     297 DSVYLGLNTDYFIETGFASKIQVSEEST-------FEE------DLQKGNRIRLRINTQK 343
           :  : |||| :|  ::: |    |        ||      ||: |  | |::
Db     290 ----ISASFDYFINSGLNAEVHVKIGDTPIRAVIPKEEIPANLRDLRNGESIHLKVKPGT 345

Qy     344 LNIFSADGSQNLIKGVNHG 362
           : ::  : | :::| |::|
Db     346 IYLYDRE-SGSIVKEVDYG 363


RESULT  12
O86831
ID   O86831      PRELIMINARY;      PRT;   397 AA.
AC   O86831;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE POLYAMINE ABC-TRANSPORTER ATP-BINDING P ROTEIN.
GN   SC8B7.10C.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Murphy L., Harris D.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR   EMBL; AL031225; CAA20221.1; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transport.
SQ   SEQUENCE   397 AA;  42352 MW;  11D6EE922D9203F9 CRC64;


  Query Match           35.9%;  Score 661.5;  DB 2;  Length 397;
  Best Local Similarity  38.3%;  Pred. No. 2.1e-36;
  Matches  142;  Conservative   76;  Mismatches  134;  Indels   19;  Gaps    6;
```

```
Qy       3 EIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEFYF 62
Db      18 DVRLAGISKTYGSFTAVHPLDLTVPQGSFFALLGASGCGKTTTLRMIAGLEEPSGGTVHL 77

Qy      63 DDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLELMQ 122
Db      78 GDQDVTALPPYKRPVNTVFQSYALFPHLDIYENVAFGLRRRGI--KSVKKQVGEMLDLVQ 135

Qy     123 IAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREIQHE 182
Db     136 LGEQARKKPHQLSGGQQQRVAVARALINHPKVLLLDEPLGALDLKLRRQMQLELKRIQTE 195

Qy     183 VGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGRTNIIPA 242
Db     196 VGITFVHVTHDQEEAMTMADQVAVMNAGRVEQLGAPADLYENPRTTFVANFLGTSNLIEA 255

Qy     243 NLEKRSDGAYIVFSDGYALRMP----ALDQVEEQAIHVSIRPEEFI---KDESGD----- 290
Db     256 EVDTRSGDDVVVKAAGDKLVLPGARCSAPAKTGGKVLVGVRPEKISLTHADDAGSIPEGR 315

Qy     291 --IEGTIRDSVYLGLNTDYFIETGFASKIQVSEESTFEED-LQKGNRIRLRINTQKLNIF 347
Db     316 NRITGRIGSTSFIGVSTQYVVDCAACPEFEVYAQNIDRDPRLTPGAEVVLHWN--PAHTF 373

Qy     348 SADGSQNLIKG 358
Db     374 GLDADQSVLAG 384
```

RESULT 13
O57933
ID O57933 PRELIMINARY; PRT; 375 AA.
AC O57933;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 375AA LONG HYPOTHETICAL MULTIPLE SUGAR-BINDING TRANSPORT ATP-BINDING
DE PROTEIN.
GN PH0194.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000001; BAA29263.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 375 AA; 42511 MW; 4DED4764FC53DA81 CRC64;

Query Match 35.9%; Score 661; DB 17; Length 375;
Best Local Similarity 38.9%; Pred. No. 2.1e-36;
Matches 147; Conservative 83; Mismatches 108; Indels 40; Gaps 12;

```
Qy       1 MSEIKIINAKKIYHDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGEF 60
Db       4 MAEVKLINIWKRFGDVTAVKDLSLEIKDGEFLVLLGPSGCGKTTTLRMIAGLEEPTRGQI 63

Qy      61 YFDDTKINNME------PSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQT 114
Db      64 YIEDNLVADPEKGVFVPPKERDVAMVFQSYALYPHMTVYDNIAFPLKLRKVPKQEIDKRV 123

Qy     115 NKYLELMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQ 174
Db     124 REVAEMLGLTELLNRKPRELSGGQRQRVALGRAIIRRPKVFLMDEPLSNLDAKLRVKMRA 183

Qy     175 AIREIQHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFI 234
Db     184 ELKKLQRQLGVTTIYVTHDQVEAMTMGDRIAVMNKGELQQVGTPDEVYYKPVNTFVAGFI 243

Qy     235 GR--TNIIPANLEKRSDGAYIVFSDGYALRM--PALDQVEEQ-----AIHVSIRPEEFIK 285
Db     244 GSPPMNFLDATI---TDDGFLDFGE-FKLKLLQDQFEVLEEENMVGKEVIFGIRPED-VH 298

Qy     286 DESG---DI--EGTIRDSVYLGLNTDYFIETGFASKI------QVSEESTF--EEDLQKG 332
Db     299 DASFTHIDVPEENTVKATVDI-------IENLGGEKIVHLRRGNISFTAKFPKESKVREG 351

Qy     333 NRIRLRINTQKLNIFSAD 350
Db     352 DEVSVVFDMKKIHIFRKD 369
```

RESULT 14
Q49978
ID Q49978 PRELIMINARY; PRT; 392 AA.
AC Q49978;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UGPC (PROBABLE ABC-TRANSPORT PROTEIN, ATP-BINDING COMPONENT).
GN ML1089.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL; U15180; AAA62878.1; -.
DR EMBL; AL583920; CAC31470.1; -.
DR Leproma; ML1089; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 392 AA; 42775 MW; 7123DF2D7C87BFDF CRC64;

Query Match 35.7%; Score 658.5; DB 16; Length 392;
Best Local Similarity 38.7%; Pred. No. 3.2e-36;

```
  Matches  153;  Conservative   57;  Mismatches  134;  Indels   51;  Gaps   10;

Qy       1 MSEIKIINAKKIY-HDVPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
           |:|| : :  | | :    : :|:||:  |    |:| ||||||| | ||||   |  ||
Db       1 MAEIVLEHVNKNYPNGATAVHDLSITVADGEFLILIGPSGCGKTTTLNMIAGLEDISSGE 60

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
             |  ::|   |  |:| ||||:||::||:||| |:|| ||  || | |: |: ::  :
Db      61 LRIDGDRVNEKAPKDRDIAMVFQSYALYPHMTVRQNIAFPLMLAKVKKAEIAQKVSETAQ 120

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :: :    |||| :|||||:||| :  |:  :|   |||||||||:||||:  |  |  :
Db     121 ILDLTDLLDRKPSQLSGGQRQRVAMGRAIVRHPKAFLMDEPLSNLDAKLRVRTRGEIARL 180

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR--T 237
           |  :| |||||||||| ||| : |:: ||: ||:|||| | ||| :| | ||| |||
Db     181 QRRLGATTVYVTHDQTEAMTLGDRVVVMRSGVVQQIGTPDELYERPVNLFVAGFIGSPVM 240

Qy     238 NIIPANLEKRSDGAYIVFSDGYALRMPALD----------QVEEQAIHVSIRPEEFIK 285
           |  || |              :|  | :  |:                | :: | | :|||
Db     241 NFFPAKLT----------PNGLTLPLGELNLSRDVQAMIGQHPVPDRVI-VGVRPEHLTD 289

Qy     286 DESGDI--EGT-----IRDSVYLGLNTD---YFIETG----------FASKIQVSEES-- 323
             :  |    ||     ::  :   |  |   ||  ||           ||:::| |
Db     290 AKLIDAHQHGTVLTFKVKVDLVESLGADKYIYFTTTGCDVYSAQLDELASELEVRENQFV 349

Qy     324 ---TFEEDLQKGNRIRLRINTQKLNIFSADGSQNL 355
               : |  :  |  | |   | |: :| ||   ||
Db     350 ARVSAESKMAIGESIELAFGTAKIAVFDADSGVNL 384


RESULT  15
Q9X103
ID   Q9X103      PRELIMINARY;      PRT;   369 AA.
AC   Q9X103;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SUGAR ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN   TM1276.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
DR   EMBL; AE001783; AAD36351.1; -.
DR   TIGR; TM1276; -.
DR   InterPro; IPR003593; AAA.
DR   InterPro; IPR003439; ABC_transportr.
DR   InterPro; IPR001687; ATP_GTP_A.
DR   Pfam; PF00005; ABC_tran; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   369 AA;  41539 MW;  AA589E22C2A56F43 CRC64;


  Query Match           35.7%;  Score 657.5;  DB 16;  Length 369;
  Best Local Similarity   38.8%;  Pred. No. 3.4e-36;
  Matches  147;  Conservative   72;  Mismatches  125;  Indels   35;  Gaps    9;

Qy       1 MSEIKIINAKKIYHD-VPVIENLNITIPKGSLFTLLGASGCGKTTLLRMIAGFNSIEGGE 59
           |::: : |  |:| : |  ::| |: :       ||| ||||||| ||||||   |  |:
Db       3 MAQVVLENVTKVYENKVVAVKNANLVVEDKEFVVLLGPSGCGKTTTLRMIAGLEEITDGK 62

Qy      60 FYFDDTKINNMEPSKRNIGMVFQNYAIFPHLTVRDNVAFGLMQKKVPKEELIQQTNKYLE 119
             | |   :|::|  |:| |||||||::|||:|| :|:|||  :| ||:|: ::  :  :
Db      63 IYIDGKVVNDVEPKDRDIAMVFQNYALYPHMTVYENMAFGLKLRKYPKDEIDRRVREAAK 122

Qy     120 LMQIAQYADRKPDKLSGGQQQRVTLACALAVNPSVLLMDEPLSNLEAKLRLDMRQAIREI 179
           :: |    |||| :|||||:||| :  |:  || | | |||||||:||||: ||  ::::
Db     123 ILGIENLLDRKPRQLSGGQRQRVAVGRAIVRNPKVFLFDEPLSNLDAKLRVQMRSELKKL 182

Qy     180 QHEVGITTVYVTHDQEEAMAISDQIAVMKDGVIQQIGRPKELYHKPANEFVATFIGR--T 237
           |  :  | :|||||| ||| ::|:| |||||  ||||| | |:|: ||| ||| |||
Db     183 HHRLQATIIYVTHDQVEAMTMADKIVVMKDGEIQQIGTPHEIYNSPANVFVAGFIGSPPM 242

Qy     238 NIIPANLEKRSDGAYIVFSDGYALRMP-----ALDQVEEQAIHVSIRPEEFIKDE----- 287
           | : | :  | :|    : : |: :::|      |    :: |   ||||: | |:
Db     243 NFVNARV-VRGEGGLWIQASGFKVKVPKEFEDKLANYIDKEIIFGIRPED-IYDKLFALA 300

Qy     288 ---SGDIEGTIRDSVYLGLNTDYFIETG----FAS---KIQVSEESTFEEDLQKGNRIRL 337
                | | :     ||  |   :: |     ||   : |  ||           :| |
Db     301 PSPENTITGVVDVVEPLGSETILHVKVGDDLIVASVNPRTQAKEE----------QKIDL 350

Qy     338 RINTQKLNIFSADGSQNLI 356
           ::  ::: |  :  : :|
Db     351 VLDMTRMHAFDKETEKAII 369


Search completed: July 30, 2002, 15:51:35
Job time: 228 sec
```